# Addressing barriers to human capital accumulation: Essays in development and health economics

Dissertation to acquire the doctoral degree from the Faculty of Economic Sciences at the Georg-August-Universität Göttingen

Submitted by

**Sophie Ochmann**
Born 23 August 1994

Göttingen, August 2023

Examination committee:

| | |
|---|---|
| First examiner: | Prof. Dr. Sebastian Vollmer |
| Second examiner: | Prof. Dr. Renate Hartwig |
| Third examiner: | Prof. Jennifer Manne-Goehler, MD, ScD |

Published as: doi:10.53846/goediss-10212

## Abstract

While health and education, jointly referred to as human capital, are important ends in themselves, they are also important drivers of poverty alleviation and economic growth. Understanding and overcoming the barriers that constrain human capital accumulation is hence crucial for economic development. This dissertation examines three barriers to human capital accumulation in three essays.

Essay one studies whether providing school-based management committees with a grant and training can improve primary educational attainment in Sokoto, Nigeria. We thereby contribute evidence from an understudied setting, a low-income context, by evaluating a large-scale intervention with a cluster-randomized controlled trial. We find that the intervention does not have any statistically significant effect on school infrastructure, student enrolment, student or teacher absenteeism or students' learning outcomes. High levels of teacher absenteeism, among others, seem to be a likely explanation for these null results.

Essay two benchmarks diagnostic testing for hypertension, diabetes and hypercholesterolaemia, three major risk factors of cardiovascular disease, against the World Health Organization Package of Essential Non-Communicable Disease Interventions testing recommendations in 57 low- and middle-income countries. We determine overall testing, its targeting according to testing need and its correlation with sociodemographic characteristics. We find adherence to testing recommendations to be low, with many individuals being tested despite not meeting the testing criteria. Additionally, the likelihood of being tested is significantly correlated with individuals' sociodemographic characteristics: women were statistically significantly more likely to be tested, as were wealthier and more educated individuals.

Essay three determines the impact of patent expiry on statin consumption in Germany, England and Sweden using the synthetic control method (SCM) which has not been used in the patent expiry literature before. We show that SCM is a suitable method for this research question and that the consumption of individual statin molecules increases upon their patent expiry by displacing other, substitute statin molecules. All three countries exhibit a high price elasticity in statin consumption, indicating a prioritization of cost-saving over the minimization of side effects.

## Zusammenfassung

Bildung und Gesundheit, zusammen auch Humankapital genannt, sind sowohl wichtige Ziele an sich als auch wichtige Faktoren für Armutsbekämpfung und Wirtschaftswachstum. Für die wirtschaftliche Entwicklung eines Landes ist es deswegen unerlässlich, Faktoren, die eine Verbesserung des Humankapitals behindern, zu erkennen und zu beseitigen. Daher untersucht diese Dissertation drei Hindernisse zur Verbesserung von Humankapital.

Der erste Aufsatz untersucht, welche Auswirkungen die Auszahlung von Geldern und Trainingsangebote für Eltern und Lehrpersonal auf die Grundschulbildung in Sokoto, Nigeria, haben. Wir präsentieren erstmalig Ergebnisse einer groß angelegten Intervention aus einer von niedrigem Einkommen geprägten Region, die wir anhand einer randomisierten Kontrollstudie evaluieren. Wir zeigen, dass die Intervention keinen statistisch signifikanten Effekt auf die Schulinfrastruktur, die Schüler*innenzahlen, die Präsenz von Schüler*innen oder Lehrenden oder die Mathematik-, Lese- und Schreibfähigkeiten der Schüler*innen hatte. Eine wahrscheinliche Erklärung für die Nullergebnisse ist die hohe Abwesenheit der Lehrenden.

Der zweite Aufsatz vergleicht das diagnostische Testen auf Bluthochdruck, Diabetes und Hypercholesterinämie, drei wichtige Risikofaktoren für Herzkreislaufkrankheiten, mit den Testrichtlinien der Weltgesundheitsorganisation (WHO) in 57 Ländern niedrigen oder mittleren Einkommens. Wir bestimmen, wie viel allgemein und gezielt entsprechend der Richtlinien der WHO getestet wird und inwieweit Testleistung mit soziodemografischen Eigenschaften korreliert. Die Einhaltung der Richtlinien ist niedrig; viele Individuen werden auf Herzkreislauf-Risikofaktoren getestet, obwohl sie die Kriterien der Testempfehlungen nicht erfüllen. Außerdem korreliert die Wahrscheinlichkeit, getestet zu werden, signifikant mit soziodemografischen Eigenschaften: Frauen, sowie wohlhabendere oder gebildetere Personen werden statistisch signifikant häufiger getestet.

Der dritte Aufsatz bestimmt anhand der synthetischen Kontrollmethode (SKM), die bislang nicht in der Patentliteratur genutzt wurde, den Einfluss von Patentausläufen auf den Konsum von Statinen in Deutschland, England und Schweden. Wir zeigen, dass SKM eine geeignete Methode für die Beantwortung dieser Forschungsfrage ist und dass der Konsum von Statinmolekülen nach Patentablauf ansteigt, indem andere Substitutsstatine verdrängt werden. In allen drei Ländern stellen wir eine hohe Preiselastizität fest, sodass die Minimierung von Kosten gegenüber der Minimierung von Nebenwirkungen priorisiert wird.

# Acknowledgments

This thesis, like most meaningful endeavors, has been the work of many hands.
To my supervision and examination committee: Thank you, Sebastian Vollmer, for staying in touch after my internship with Stephan Klasen in 2015 and encouraging me to return to Göttingen for my doctoral studies. I am grateful to you for supervising me, for offering research opportunities with the NIPEP team in Sokoto, Nigeria, and part of the HPACC team at Harvard University, and for trusting me with various teaching tasks at the undergraduate and graduate level and the setting up of a new Bachelor program in Sustainable Development Studies. Thank you, Renate Hartwig, for your helpful advice and comments, especially on my third essay. I am sorry that I did not manage to engage more with you in the last four years. Thank you, Jennifer Manne-Goehler, for welcoming me so warmly in the HPACC team and always, without exception, being such an upbeat and driven supervisor so ready to share your medical wisdom or experience with medical journal publications with me. Please keep up your contagious positivity.

To all other co-authors: Thank you for sharing the road of this PhD journey with me and thereby making this final product possible. I would like to especially thank Isabelle von Polenz for being a great collaborator and friend and Jessy Amarachi Ezebuihe for being an excellent research assistant and partner for the data collection in Sokoto. Thank you also to S Anukriti and Catalina Herrera Almanza for being such passionate researchers I have been honored to collaborate with and learn so much from. Even though our research on female vocational students' school-to-work transition in Haryana, India, did not make it into my dissertation, I drew a lot of energy from our research discussions. I hope I will get to meet you both face-to-face one day.

To the chair of Sebastian Vollmer, including those that have moved on to different academic homes by now: Thank you for providing a fun and collaborative work environment 'without elbows'. Thank you to Amal Ahmad, Dominik Naeher and Lisa Bogler for reading earlier versions of parts of this dissertation. I would like to give a special shout-out to Ann-Charline Weber and Lisa Bogler who form the backbone of this chair: thank you for taking care of

countless active break registrations, birthday, wedding and newborn presents, and welcoming every new chair member so warmly.

To my wonderful friends, who truly shaped the past four years into a memorable and joyful journey (in no particular order): To Annkathrin, my fellow PhD friend and, together with Ghiath, my one COVID contact household: thank you for the countless shared meals, heart-to-heart conversations and making me feel like your door is always open for me. To Sarah and Milo: thank you for always being a home away from home for me and making our friendship endure all the distance and time zone differences. To Paula and Esther: thank you for the trips, kilos of asparagus and wisdom that you shared with me. To Adi, thank you for our fun conversations about tax policy, dogs and Western Cape and also for frequently letting me third wheel you and Philipp. To Tebello, thank you for being an inspiration for self-love and the pursuit of one's passions, and for introducing me to Zadie Smith's novels. To all the friends that I acquired through our shared love for cycling, with whom I loved riding, racing and sharing a pizza. Thank you, Kamila, Lisa, Pascal, Kay, Evi, Lina and the whole Tuspo Weende crew for the (type II) fun memories together. To the medicine girls: Thank you Christa, Summer, Miri, Rox, Becci, Mandy and Caro for welcoming me so warmly as an economist and explaining details necessary for the medical part of this thesis to me. Thank you also to my dear friends Amma, Gautam, Dana, Hannah, Hanna, Ruby, Vanessa, Victoria and Vidushi for the visits, trips and friendships we share.

To my family, Susanne, Thomas, Ariane, Moritz, Peter, Christa, Tini, Ulf, Anna Marie, Charlotte, Florian: thank you for your unfaltering love and support throughout my life which got me to the start line of this PhD in the first place and for unwaveringly being by my side throughout it. Celebrating my successful defense with the kissing of the Gänseliesel as a Göttingerin born and bred really was a special day that I will never forget and am glad to have shared with you.

Last but not least, I want to thank my partner Nils for believing in me and sharing my passion for cycling and Golden Retrievers. Thank you for keeping me sane with your calm and assured presence and for showing me that there is more to life than professional achievements.

Sophie Ochmann
Göttingen, November 2023

# Contents

## Index of figures in main text



## Index of tables in main text

# Index of figures in the appendix

# Index of tables in the appendix

# 1. Introduction

## 1.1. Background and motivation

Good health and education are an important end in itself: They are capabilities that empower people to lead the free and fulfilling lives they have reason to value (Sen, 1997). Economists augment this view by also considering health and education as important means for generating economic growth. For instance, already Adam Smith thought about the role of education in production processes as a division of labor was only possible with learning new skills and specialization (Smith, 1776). That is why health and education are often referred to as human capital, a wording that reflects their role as input factors into production functions alongside physical labor and capital.[1]

Human capital additionally plays an important role for poverty alleviation and economic development as healthier and more educated individuals can work longer and more productively and complete tasks requiring more advanced skillsets (World Bank, 1993; 2017). For instance, men who were exposed to a deworming program as primary school children were able to work more hours and missed fewer meals per week (Baird et al., 2016). Conversely, persistent malaria infections during childhood can reduce adult income by 50 percent (Bleakley, 2010). And an additional year of schooling is estimated to yield a 10% rate of return, approximately (Duflo, 2001; Montenegro and Patrinos, 2014).

While there have been impressive gains in health and education globally in recent decades, many more improvements are needed before everyone can freely access high quality education and health services irrespective of where they are born. Child mortality (the number of deaths under the age of five per 1,000 live births) has decreased from 9.32 in 1990 to 3.66 in 2020 while global literacy has increased from 74.6% to 86.8% (World Bank, 2023a).

[1] Many macroeconomic models of economic growth include human capital, for instance the Augmented Solow Model (Mankiw, Romer and Weil, 1992). Empirically, when including human capital in growth regressions, Mankiw, Romer and Weil also find a better empirical fit of their model, ie they are better able to explain cross-country differences in national income in comparison to a model only including physical capital and labor (1992).

At the same time, we are unlikely to reach the Sustainable Development Goals by 2023 (Our World in Data, 2023), which call for an end of all preventable deaths under five years of age and universal literacy, along with even more ambitious targets like "achieve universal health coverage […] for all" or universal completion of "free, equitable and quality primary and secondary education" (United Nations, 2023) by 2030.

Additionally, researchers and policy makers have expanded the list of health and education issues that urgently should be understood and improved.

For education, the focus has shifted away from improving educational attainment like enrolment or completion rates towards prioritizing learning. Globally, children spent an average of 11.2 years in school but obtained a learning achievement corresponding to only 7.9 years in school in 2018 (World Bank, 2020). The literature even termed this the 'global learning crisis' (World Bank, 2017). Additional shifts in the education literature have been dedicated to making interventions cost-effective and scalable (Angrist et al., 2020; Bold et al., 2018; Angrist and Meager, 2023).

For health, in turn, researchers and policy makers have had to expand the scope of their work as the disease burden in the Global South is now characterized by a so-called double burden of disease: Low- and middle-income countries (LMICs) face both infectious diseases like HIV/AIDS or malaria and non-communicable diseases like cardiovascular disease (CVD) (WHO, 2022). CVD describes conditions relating to the heart or blood vessels like strokes or heart attacks (NHS, 2023a) and is the number one cause of death in LMICs (15 million out of a total of 46 million deaths in 2019, IHME, 2023).

## 1.2. Literature and contribution

This thesis contributes to understanding and addressing these shifts. To the education literature, essay 1 adds evidence on the impact of a large-scale primary school intervention on students' learning outcomes. Essays 2 and 3 are related to health and provide new insights into the testing for and preventing of cardiovascular disease.

The education literature provides a plethora of evidence on various input factors necessary for improved schooling and learning outcomes (review by Glewwe and Muralidharan, 2016), ranging from supply-side input factors like the number (Angrist and Lavy, 1999), presence (Duflo, Hanna and Ryan, 2012) and effort (Muralidharan and Sundararaman, 2011) of teachers or the number of schools (Duflo, 2001) to demand-side input factors like children's health (Hamory et al., 2021), parents' beliefs (Benhassine et al., 2015) or household income (Baird, McIntosh and Özler, 2011). Two important determinants of successful education interventions that this literature gives rise to are timing and consideration of local circumstances.

Acquiring basic numeracy and literacy skills early is crucial. Children learn these skills easiest at primary school age when their brain is still malleable (Cunha et al., 2006).[2] Literate and numerate children can master more advanced lessons and are thereby able to progress with schooling and also have better labor market outcomes later in life (Pritchett and Beatty, 2012; World Bank, 2017). Having basic literacy and numeracy skills has benefits for other, non-labor market related outcomes as well, such as individuals' health behavior (de Walque, 2007), fertility decisions (Güneş, 2016), investment in children (Chen and Li, 2009) or civic engagement (Larreguy and Marshall, 2017).

Many well-intentioned education programs have no or very little impact on schooling or learning because the local context is ignored. Numerous examples exist where programs did not address the binding constraint (the 'bottleneck') that inhibited educational attainment or learning. For example, distributing textbooks in Kenya only benefitted the best performing students as the English language of the textbooks created a language barrier for most students (Glewwe, Kremer and Moulin, 2009). Similarly, providing Peruvian primary school children with a free laptop for home use had no impact on their learning outcomes and even deteriorated their effort as reported by teachers (Beuermann et al., 2015).

To target interventions better at the true bottlenecks of children's education, school-based management committee (SBM or SBMC) interventions were born. SBMCs are usually

[2] In fact, there are studies that argue that the time before starting with primary school is also very important for children's cognitive development and their life outcomes as an adult (Almond, Mazumder and Van Ewijk, 2015; Gertler et al., 2014).

comprised of parents and teachers (sometimes also a student representative) and are supposed to be the decision-makers and monitors of improvements regarding the local school, decentralizing the authority from the central government to the school level (Barrera-Osorio et al., 2009). In this approach, parents and teachers are thought to be suited best for determining the binding constraints to children's education, with parents having a clear incentive to improve their own children's education. SBMC interventions therefore usually give committee members trainings on how to function as monitors and change-makers or provide grants to finance desired improvements. Evidence on the success of these programs, however, has so far been mixed at best and is usually derived from studies of small pilot programs (e.g. Lassibille, Tan, Jesse and Van Nguyen, 2010; Blimpo, Evans and Lahire, 2015; Banerjee et al., 2010). One exception form two studies from Mexcio, an upper middle-income context (Garcia-Moreno, Gertler and Patrinos, 2019; Santibañez, Abreu-Lastra and O'Donoghue, 2014).

Essay 1 of this dissertation provides evidence on the state of primary education in a new, low-income context and evaluates the impact of a large-scale SBMC intervention on primary schools' functioning and children's learning.

Turning to the second component of human capital, health:
Despite cardiovascular disease being the number one cause of death in LMICs, the state of their CVD care remains underexplored. It was only in the past five years that the 'Global Health and Population Project on Access to Care for Cardiometabolic Diseases (HPACC)' provided the first evidence based on individual-level survey data from a large sample of low- and middle-income countries on the state of care for CVD and its risk factors (eg Geldsetzer et al., 2019; Manne-Goehler et al., 2019; Flood et al., 2021; Marcus et al., 2021; Peiris et al., 2021). These studies quantified how many individuals with a certain CVD risk factor were tested, diagnosed, treated and had the risk factor under control. They found that the largest gap in this four-step care continuum for three major CVD risk factors – hypertension, diabetes and hypercholesterolaemia – was the first step: diagnostic testing. Only 74% of individuals with hypertension had ever had their blood pressure measured, 63% of those with diabetes had ever had their blood sugar measured and 43-47% of those with hypercholesterolaemia had ever had their cholesterol measured (Geldsetzer et al., 2019; Manne-Goehler et al., 2019;

Marcus et al., 2021). Diagnostic testing is crucial for detecting those with hypertension, diabetes or hypercholesterolaemia to ensure entrance into the remaining care continuum (diagnosis, treatment, control) and thereby lower their CVD risk.

However, individuals that do not have one of these CVD risk factors also undergo diagnostic testing. Exclusively considering testing among those with a CVD risk factor therefore paints only half the picture of the state of diagnostic testing in LMICs. Instead, to fully assess LMICs' diagnostic testing performance, one should consider the extent to which health systems are able to target their diagnostic testing efforts at those patients with a high risk of developing CVD.[3] In other words, an assessment of countries' CVD risk factor testing performance should evaluate whether those individuals who are being tested are those individuals with a high testing need, ie are at a high risk of developing CVD.

So far, evidence of LMICs' CVD risk factor testing of their whole population and the extent of their targeting remains scarce. The literature is based on evidence from individual countries (Nambiar et al., 2020; Ciancio et al., 2021), a study using facility-level data from only 10 LMICs (Yadav et al., 2021) and a study considering only behavioral CVD risk factors in 6 LMICs (Ruan et al., 2018).

Essay 2 of this dissertation therefore contributes the first evidence on the diagnostic testing capacity of a large sample of LMICs for three major cardiometabolic CVD risk factors using individual-level data from a large set of countries and benchmarks this against testing need.

To successfully prevent CVD, however, patients need to complete all four steps of the care continuum and have their risk factor conditions like hypertension, diabetes or hypercholesterolaemia under control. This includes access to treatments such as a statin regiment. Statins are commonly prescribed drugs to lower cholesterol and control hypercholesterolaemia as elevated cholesterol levels facilitate plaque build-up in the arteries

[3] To have a 100% success rate of catching all patients with an elevated CVD risk, countries could also pursue a universal testing strategy, regularly testing its whole population for CVD risk factors. This strategy, however, would not be feasible in real-world settings as diagnostic testing capacities are not endless, especially not in low- and middle-income country contexts.

(CDC, 2023). If plaque build-up eventually blocks an artery completely, this results in a cardiovascular disease event such as a heart attack (blocked coronary artery) or ischemic stroke (blocked blood flow to the brain) (NHS, 2023a). Access to statins is low in LMICs as only one in ten hyperlipidaemic individuals received statins for primary CVD prevention (Marcus et al., 2022).[4]

Two important determinants of low access to treatment of CVD risk factors in the Global South are a low health insurance coverage and high costs. Health insurance coverage remains low, with average coverage being 7.9% in low-income countries, 27.3% in lower-middle-income countries, and 52.5% in upper middle-income countries (Hooley et al., 2022). As a consequence, individuals are forced to pay for health services and treatments out of pocket and spend large portions of their income on health (Dupas, 2011). Xu et al. (2003) find the proportion of households that had to spend more than 40% of their net income to be more than 10% in Brazil and Vietnam. However, the price elasticity of individuals in LMICs to treatment costs differs by the treatment type: Demand for curative treatment is usually found to be quite price inelastic (Cohen, Dupas and Schaner, 2015) while demand for preventative health measures is usually quite price elastic which means that they react to price increases with a substantial reduction in demand (Cohen and Dupas, 2010). This difference in elasticity is therefore likely to exacerbate the impact of treatment cost acting as a barrier to preventative treatment coverage.

Understanding the factors that drive the price of CVD preventing treatments like statins is important as it is a key predictor of treatment uptake in the Global South. Unfortunately, this research is hindered by a scarcity of data on drug prices and especially drug consumption in low- and middle-income countries. In high-income contexts, the literature has been able to show that patent protections are an important driver of drug prices (review by Vondeling et al., 2018) but the impact on drug consumption remains less well understood. Only seven studies have so far looked at the impact of patent expiry on overall molecule consumption

[4] Primary prevention refers to cases where patients have never had a CVD event before. Secondary prevention, in turn, describes the steps taken to prevent additional CVD events in patients that already have a CVD event history. Marcus et al. (2022) find that in the latter case, still only one in five patients consume statins as secondary prevention.

(Aitken et al., 2013; Berndt, Kyle and Ling, 2003; Chapman, Fitzpatrick and Aladul, 2017; Duflos and Lichtenberg, 2012; Fiorentini, Bruni and Mammi, 2022; Imai, Fushimi and Sundell, 2018; Lakdawalla and Philipson, 2012)[5] which predominantly examined the United States and used descriptive or regression-based analyses rather than rigorous causal identification techniques.

Essay 3 contributes causal evidence of the impact of patent expiry on drug consumption from several high-income countries in Europe to this literature, using statins as the example drug class.

## 1.3. Chapter overview

In the following, I summarize the three essays of this dissertation.

**Essay 1: The impact of grants in combination with school-based management trainings on primary education: a cluster-randomized trial in Northern Nigeria**

*Joint work with Kehinde Elijah Owolabi, Folake Olatunji-David, Niyi Okunlola and Sebastian Vollmer.*
*Published in the Journal of Development Effectiveness.*

Essay 1 examines whether an intervention that provides school-based management committees with a grant and a training on the planning of school improvements can improve school infrastructure and educational attainment and learning in Sokoto state, Nigeria.

We evaluate two components of a large-scale SBMC intervention by running a randomized controlled trial with 128 primary schools in rural and peri-urban Sokoto, a state in north western Nigeria. The SBMC intervention was called 'Nigerian Partnership for Education Project' (NIPEP) and conducted for 100 million USD in five northern states of Nigeria, funded

[5] This is because most studies instead consider the market share of brand product vis-à-vis generic producers to describe the competitive nature of pharmaceutical markets (eg Fischer and Stargardt, 2016).

by the Global Partnership of Education and the World Bank. In July and August 2018, we collected observational data on the school infrastructure and survey data from the headmaster, teachers, parents and 5,717 Grade 2 and 3 students. We also tested students' numeracy and literacy in Hausa, the local language. We randomized half of the schools (n=64) into a treatment group where SBMC members received a training and a grant by the Sokoto ministry of education. Within the group of treatment schools, we additionally introduced variation in the grant amount, half of the treatment schools (n= 32) receiving NGN 250,000 (approx. PPP-adjusted int-$ 2,250) and the other half receiving twice that amount, so NGN 500,000. We collected observational and survey data again 14 months later in November and December 2019, this time including students from Grades 2, 3 and 4 in our data collection.

We find that the intervention of training and a grant disbursal to SBMCs had no statistically significant effect on any of our outcomes, ie school infrastructure, student enrolment, teacher or student attendance, or learning outcomes. These null results also remain when disaggregating treatment effects by the grant amount that treatment schools received.

We explore five potential explanations for the intervention having no discernable impact on the learning environment or achievement of the treatment schools. One plausible factor is a faulty implementation of the intervention as we find low levels of self-reported receipt of the intervention by SBMC members as well as no improvements in the areas that SBMC members claimed to have spent the grant money on (conditional on reporting an intervention). More importantly, though, we find very high levels of teacher absenteeism with no learning taking place at 74% and no teacher being present at 45% of the schools in our sample.

In conclusion, this essay provides evidence of a large-scale SBMC intervention that failed to produce any significant results. It is an example where a well-intentioned intervention did not target the true bottleneck of primary education in this context (teacher absenteeism). While having a school building with a roof, tables, chairs and a blackboard is conducive to learning, it will not take place if no teacher is present. Future education interventions should therefore be designed by first becoming thoroughly familiar with the local context, understanding all the barriers faced by primary school children, and then determining and targeting the

bottleneck, ie the binding constraint hindering the improvement of educational attainment and learning.

**Essay 2: Diagnostic testing for hypertension, diabetes and hypercholesterolaemia in low-income and middle-income countries: a cross-sectional study of data from 994 185 individuals from 57 nationally representative surveys**

*Joint work with Isabelle von Polenz, Maja-Emilia Marcus, Michaela Theilmann, David Flood, Kokou Agoudavi, Krishna Kumar Aryal, Silver Bahendeka, Brice Bicaba, Pascal Bovet, Luisa Campos Caldeira Brant, Deborah Carvalho Malta, Albertino Damasceno, Farshad Farzadfar, Gladwell Gathecha, Ali Ghanbari, Mongal Gurung, David Guwatudde, Corine Houehanou, Dismand Houinato, Nahla Hwalla, Jutta Adelin Jorgensen, Khem B Karki, Nuno Lunet, Joao Martins, Mary Mayige, Sahar Saeedi Moghaddam, Omar Mwalim, Kibachio Joseph Mwangi, Bolormaa Norov, Sarah Quesnel-Crooks, Negar Rezaei, Abla M Sibai, Lela Sturua, Lindiwe Tsabedze, Roy Wong-McClure, Justine Davies, Pascal Geldsetzer, Till Bärnighausen, Rifat Atun, Jennifer Manne-Goehler, and Sebastian Vollmer*
*Published in the Lancet Global Health.*

Diagnostic testing for three major CVD risk factors, hypertension, diabetes and hypercholesterolaemia, is the crucial first step to effective, efficient and timely management of these risk factors and thereby cardiovascular disease risk. We quantify the diagnostic testing performance of low- and middle-income countries' healthcare systems. We compare in how far diagnostic testing status overlaps with the WHO PEN recommendations of who should be tested, ie anyone with symptoms, a BMI larger than 30 or aged at least 40 years with a BMI over 25. We additionally disaggregate and compare this diagnostic testing performance by three sociodemographic characteristics: sex, wealth and education.

We pool individual-level data from 994 185 non-pregnant adults aged at least 18 years old from nationally representative surveys done between 2010 and 2019. Countries had to be a low- or middle-income country according to the World Bank definition (World Bank, 2023b) at the time of the survey and had to include a question on whether respondents had ever had their blood pressure, glucose, or cholesterol measured. We create a separate analysis sample for each CVD risk factor as not all surveys included all three questions on the diagnostic testing for each risk factor. We report four key outcomes; the shares of tested individuals and individuals who met the WHO PEN criteria, the share of tested individuals among all who met

the WHO PEN criteria, and the share of the sample for whom testing guidelines were adhered to, in other words the share of those who did not meet the WHO PEN criteria and were not tested plus those who met the WHO PEN criteria and were tested out of the whole sample.

We find adherence to testing guidelines to be low – 72.2% for diabetes, 70.6% for hypercholesterolaemia and 49.0% for hypertension. This is driven by substantial shares of individuals being tested despite not meeting the WHO PEN testing criteria. At the same time, we find substantial shares to not have their blood sugar (11.5%) or cholesterol (13.5%) tested despite meeting the WHO PEN testing criteria.

Additionally, we find individuals that are female, or in higher wealth or education categories are more likely to meet the WHO PEN criteria. When controlling for these differences in testing need, however, we find that sociodemographic characteristics still significantly correlate with diagnostic testing performance. When considering only those that met the WHO PEN testing criteria, women are more likely to be tested for hypertension in comparison to men. More educated or wealthy individuals are more likely to be tested for all three CVD risk factors than those in the lowest education category or wealth quintile.

The combination of low adherence to the WHO PEN diagnostic testing criteria and sociodemographic inequalities in access to testing leaves ample room for improved targeting of CVD risk factor diagnostic testing efforts. Policy makers should ensure that all individuals who meet the WHO PEN testing criteria are tested, independently of their sociodemographic status, as access to diagnostic testing determines the outcomes further along the care continuum, ie how well these risk factors can be brought under control to thereby also prevent CVD.

**Essay 3: The impact of patent expiry on statin consumption: A synthetic control analysis**

*Joint work with Gabriele Gradl, Martin Schulz and Sebastian Vollmer.*

Patents grant pharmaceutical producers monopoly selling rights of their drug, usually for twenty years. While the impact of patent expiry on drug prices has been studied extensively

already (review by Vondeling et al., 2018), the impact on drug consumption remains less well understood. Previous studies have predominantly examined the United States and employed descriptive or regression-based methodologies.

We determine the impact of patent expiry on the consumption of simvastatin and atorvastatin, two majorly consumed statins that are reasonably good therapeutic substitutes (Weng et al., 2010). We use monthly administrative data from Germany, England and Sweden. We estimate the causal effects employing the synthetic control method, which generates a control group using a weighted average of drugs acting on the cardiovascular system but not a lipid-lowering agent. This quasi-experimental method has not been used to examine the impact of patent expiries before. We additionally employ the previously applied Interrupted Time Series method as a robustness check.

For both simvastatin and atorvastatin, we find that the patent expiry of a molecule has two effects. First, it increases the consumption of the molecule whose patent expired and, second, it decreases the consumption of other statins. We argue that the most likely channel through which patent expiry impacts drug consumption is price: The price competition by generic producers upon a molecule's patent expiry decreases its price and leads healthcare providers to substitute towards the cheaper statin. The price sensitivity of the German, English and Swedish healthcare system as observed through the high elasticity of substitution in response to the patent expiry of simvastatin and atorvastatin has resulted in major cost savings. For example, switching to generic simvastatin in the year of its patent expiry saved the German health system €220 million (Klose and Schwabe, 2004b). However, the predominant consumption of simvastatin between May 2003 (simvastatin's patent expiry) and May 2012 (atorvastatin's patent expiry) meant a heightened side effect risk and potentially also worse compliance or more discontinuation.

We conclude that a molecule's consumption increases in response to its patent expiry and that the German, English and Swedish healthcare systems prioritize cost saving over minimizing side effects. Additionally, future research efforts should be directed towards replicating these results in other contexts with other drug classes, understanding the role of

context-specific health policies such as pricing or reimbursement rules, and, last but not least, examining the impact of drug's patent expiry on health outcomes.

## 1.4. General summary and conclusion

This dissertation concludes that many challenges remain in improving health and education outcomes in the Global South. The barriers to human capital addressed in this dissertation all have a specific window of opportunity in which they should be addressed. If children in northern Nigeria do not learn how to read, write and perform simple calculations in primary school, they are unlikely to benefit from further years of schooling or reap the labor market benefits like higher incomes or more stable employment that come with a formal education. Similarly, if individuals at a high risk of experiencing cardiovascular disease events are not identified or treated early enough, either because they are not tested or do not have access to treatment due to prohibitively high costs, they are more likely to develop cardiovascular disease which could incur high curative treatment costs or, even worse, premature mortality.

Additionally, this thesis highlights the importance of targeting for health and education interventions; targeting the individuals with the highest need as well as identifying and targeting the bottlenecks that truly inhibit health and education gains. For instance, it should be the first step to determine whether teacher absenteeism or a lack of training and funding of the SBMC are what holds back primary educational attainment, or whether the price and availability of generics determines preventative treatment take-up or not.

Policy-makers and researchers should continue to strive for improving health and education globally, until every child learns how to read and write in primary school and every individual has access to universal health coverage, as the international community agreed on with the passing of the Sustainable Development Goals in 2015. To make progress towards these goals, the understanding of the true bottlenecks and the correct targeting of interventions will be key.

# ESSAY ONE

# 2. The impact of grants in combination with school-based management trainings on primary education: a cluster-randomized trial in Northern Nigeria

Joint work with:

Kehinde Elijah Owolabi, Folake Olatunji-David, Niyi Okunlola, Sebastian Vollmer

Published in the Journal of Development Effectiveness, 2022:

Ochmann, S., Owolabi, K.E., Olatunji-David, F., Okunlola, N. and Vollmer, S., 2022. The impact of grants in combination with school-based management trainings on primary education: a cluster-randomized trial in Northern Nigeria. *Journal of Development Effectiveness*, *14*(3), pp.189-208.

## 2.1.Abstract

Grant disbursals and school-based management interventions have received growing attention from policy-makers despite their mixed success at improving educational outcomes. This paper reports results from a large-scale, cluster randomized controlled trial in Sokoto state, Nigeria. School-based management committees received a training and a grant to improve access to and quality of primary school education, especially for girls. One year after implementation, the intervention had no impact on schools' infrastructure, educational attainment or learning outcomes. Therefore, understanding the context-specific constraints to primary school education is important to avoid spending 100 million USD on a program with no discernable impact.

## 2.2. Introduction

Education has featured strongly in development efforts in the past decades, with particular attention to increase access to basic education and, more recently, improving its quality to ensure learning. 59 million children of primary school age remained out of school in 2018 (UNESCO, 2019) and on average, children spend 11.2 years in school. Yet, their learning achievements only correspond to 7.9 years in school, so children spend on average 3.3 years in school without learning (World Bank, 2020). Determining what works in improving the delivery of education has therefore received growing attention among practitioners and academics but remains a highly context-specific issue. Two prominent approaches have produced mixed results at improving the supply of quality education: First, providing schools with grants to improve their learning environment, eg infrastructure or working materials, thereby utilizing communities' insights into which local conditions act as binding constraints to educational attainment, has not proven effective (Newman et al., 2002; Olken, Onishi and Wong, 2014; Das et al., 2013). Second, empowering local school communities by providing management trainings has worked in some contexts (Lassibille et al., 2010) under certain conditions (Pradhan and De Ree, 2014; Blimpo, Evans and Lahire, 2015) but not always (Banerjee et al., 2010; Santibañez, Abreu-Lastra and O'Donoghue, 2014; Glewwe and Maïga 2011; Pradhan and De Ree 2014).[6] However, only two of these evaluations (both in Mexico) concerned large-scale interventions involving more than 10,000 primary schools (Garcia-Moreno, Gertler and Patrinos, 2019; Santibañez, Abreu-Lastra, and O'Donoghue, 2014).

Here, we assess the joint impact of a grant disbursal plus training program for school-based management committees of Nigerian primary schools. We evaluate the hypothesis that the combination of empowering local communities in identifying constraints in the supply of quality primary education as well as providing the financial means to alleviate these will improve the educational attainment as well as learning outcomes of primary school students. We do this in a rural, high- poverty setting with a poorly functioning primary school system.

[6] For a good, if slightly outdated, overview see Barrera-Osorio, Fasih and Patrinos (2009).

This paper uses a large field experiment with 128 primary schools in rural Nigeria. Half the schools were randomly selected into a treatment where each school-based management committee (SBMC) received a leadership and school management training as well as a school improvement grant. Half of the treatment schools (n = 32) received the normal amount of 250,000 NGN (approx. PPP-adjusted int-$ 2,250) as per NIPEP guidelines while the other half received twice that amount, so 500,000 NGN. The normal grant amount in the rural Sokoto context is enough to pay 10 qualified teachers their entry-level salary for a year, provide 120 students with school uniforms or construct two toilet buildings. Our analysis is based on surveys with headmasters, teachers, SBMC members and 6,000 primary school students. We also tested students' literacy and numeracy skills.

We find that the intervention had no discernible impact on schools' infrastructure or equipment, enrolment, student or teacher attendance and learning achievements, regardless of high or normal grant amount treatment status. Anecdotally, some schools that had no toilets prior to the intervention seem to have used the grant money to build some toilets. We postulate five potential reasons for the zero result.

First, challenges in the implementation may have meant that schools never received any grants, which is potentially corroborated by the low reporting of the intervention – only 50% of SBMC members and headmasters at treatment schools reported receiving an intervention at endline.

Second, schools may have received the grants but then decided to use the grants for school-unrelated matters, which is anecdotally supported by our data showing no improvements in the areas that respondents claimed to have spent the grant money on.

Third, schools' infrastructure was so wanting that the grant amount could be insufficient to alleviate this binding constraint as an input factor into the schooling production function. Even though 83% of sample schools had some sort of permanent structure, these were often in dire condition and/or used for other purposes such as storing harvests.

Fourth, educational attainment may not have been limited by the learning environment realities but by teachers' absenteeism. At 45% of schools, there was no teacher present upon the arrival of enumerators as part of the endline survey, and at 74% no learning was taking

place.[7] Improving the school-based management committee's managerial and financial capacity may therefore not have addressed the core issue that constrains educational attainment in the study context.

Fifth, SBMC's capacity may be insufficient for the training and grant disbursal to be converted into primary schools' improvements. Following Blimpo, Evans and Lahire (2015), we estimate heterogeneous treatment effects of the intervention by the literacy rate of schools' SBMCs. We find no differential treatment effect by SBMC capacity, potentially due to the low baseline capacity with SBMCs' mean literacy being at 44%.

This study contributes to the ongoing debate in the education literature (Global Education Evidence Advisory Panel, 2020) on the (cost-)effectiveness of increasing access to traditional inputs such as schools, buildings, textbooks or uniforms vis-à-vis pedagogical interventions such as Teaching At the Right Level (Banerjee et al., 2007). When reviewing education interventions' (cost-) effectiveness, Angrist et al. (2020) find that school-based management trainings in combination with grants have a mean zero impact on learning with few positive-impact outliers. Khattri, Ling and Jha (2012) and Yamauchi (2014) observe positive impacts of school-based management (SBM) reforms in the Philippines, and Gertler, Patrinos and Rubio-Codina (2012) detect reduced grade failure and grade repetition in response to an SBM reform in rural Mexico, though the positive impact vanishes in extremely poor communities. Beasley and Huillery (2017) observe improvements in enrolment and schools' resources in Niger but, with a simultaneous increase in teacher absenteeism, detect no impact on learning outcomes. Blimpo, Evans and Lahire (2015) observe a reduction in student and teacher absenteeism but no impact on student test scores.

This paper makes an important contribution to understanding what does not work in education policy. The intervention evaluated was part of the larger 'Nigerian Partnership for Education Project' (NIPEP), a 100 million USD program funded by the Global Partnership of Education and the World Bank. From 2015 to 2020, more than 28,000 primary schools in five states in Northern Nigeria received School Improvement Grants, so that – excluding

[7] Consequently, at 29% of schools, teachers were present but not teaching.

administrative costs – the primary school grant component of NIPEP alone already cost approximately 7 billion NGN or 21.6 million USD[8]; money that could have achieved substantial learning achievements if spent on projects that deliver impact in a cost-effective way (e.g. Kenya's national literacy program Tusome (Piper et al., 2018), see Angrist et al. (2020) for a review).

The remainder of this paper proceeds as follows: Section 2 describes the setting, sample, experimental design, data and estimation strategies while Section 3 presents the main results. Potential reasons for our null results are outlined in Section 4 before we conclude in Section 5.

## 2.3. Experimental design and data collection

### Study setting

The study took place in nine rural and peri-urban Local Government Areas (LGAs) in Sokoto State in the north-west of Nigeria. Sokoto state is the state with the highest poverty headcount rate, with 87.7% of the population living on less than $1,90 a day at 2011 PPP international prices in 2019 (Nigerian Bureau of Statistics, 2020).

Sokoto performs similarly on socio-demographic dimensions in comparison to nationwide averages. Only 40.9% of Sokoto's population is literate (Nigerian National Bureau of Statistics, 2020) and of all children aged 5-16 years only 21.8% were literate and 10.6% were numerate (National Population Commission, 2016). Primary school enrolment is less than 60% and primary school attendance only amounts to 40.4%. Fertility is still high with women bearing on average 7.3 children and only 40.1% of households have access to electricity (see Table 1).

[8] The School Improvement Grants were delivered to 28,049 primary schools (World Bank, 2021). Administering 250,000 NGN per school yields a total amount of grant money disbursed of 7.012 billion NGN. Using the 2019 World Bank's alternative conversion factor of 325,0 results in a total grant amount disbursed to primary schools of 21.576 million USD.

## Sampling

The sample of 128 primary schools was constructed by selecting nine Local Government Areas (LGAs) of Sokoto state where (a) the program had not yet been implemented and (b) the security situation in June 2018 was deemed safe enough for surveys to take place. The nine LGAs were Binji, Bodinga, Goronyo, Ilela, Kware, Silame, Taumbuwal, Wamakko and Wurno. Schools included in the sampling frame had to be eligible for NIPEP[9] and had to have between 35 and 160 registered Grade 2 students.

*Table 1: Sociodemographic comparison of Sokoto state and Nigeria*

| | Sokoto | Nigeria |
|---|---|---|
| Average household size | 5.93 | 5.06 |
| Fertility rate | 7.3 | 5.8 |
| Literacy in any language (in %) | 40.9 | 63.2 |
| Gross primary school enrolment rate (% of school age population) | 59.6 | 88.6 |
| Net primary school attendance (% school age population) | 40.4 | 65.8 |
| Access to electricity (% total number of households) | 40.1 | 63.7 |

Note: Fertility rate data taken from Nigerian National Bureau of Statistics and UNICEF (2017); remaining data taken from Nigerian National Bureau of Statistics (2020).

Surveys were conducted with the headmaster, additional teachers, available SBMC members and up to 25 Grade 2 and 25 Grade 3 pupils – as part of the endline survey, Grade 4 pupils were also interviewed.[10] Additionally, surveyed pupils were also given short tests in mathematics and Hausa, the local language.

## The intervention: NIPEP

The Nigerian Partnership for Education Project (NIPEP) was funded by the Global Partnership for Education, developed by the World Bank and implemented by the Federal Ministry of

[9] Primary schools were eligible for receiving a School Improvement Grant (SIG) if they had (i) a functioning SBMC, (ii) received SBMC training in the administration of SIGs, (iii) a School Improvement Plan (SIP), and (iv) established a functioning bank account (World Bank, 2015).

[10] In most schools, less students were enrolled or present on the day of the survey, so that simply all Grade 2 and 3 students were interviewed. If more than 25 students were available per grade, at baseline a random sample was supposed to have been drawn but field observations showed these were unsuccessful. Therefore, at endline, a convenience sampling methodology was officially adopted.

Education in five states in Northern Nigeria. Its aim was to 'improve access and quality of basic education [. . .], with particular attention to girls' participation.' (World Bank, 2015). The entire program consisted of three major components:

Component 1: Promoting School Effectiveness and Improved Learning Outcomes

(a) School Improvement Grants to Primary Schools
(b) School Improvement Grants to Pre-Primary Schools
(c) Support to Teachers' Professional Development

Component 2: Increasing Access to Basic Education for Out-of-School Children with focus on Girls

(a) Scholarships for Girls
(b) Scholarships for Female Teachers
(c) Community Mobilisation and SBMC Training

Component 3: Strengthening Planning and Management Systems including Learning Assessment and Capacity Development

a) Management and Implementation Support (for the Federal Ministry of Education)
b) Monitoring, Evaluation and Learning Assessment

*Grant and SBMC training components*

The evaluation of the intervention presented here focused on Components 1(a) and 2(c) above, so that schools in the sample only received School Improvement Grants and the training program for the School-Based Management Committee (SBMC). The School Improvement Grants (SIG) amounted to 250,000 Nigerian Naira (approx. PPP-adjusted int-$ 2,250) and were intended for 'non-salary expenditures related to improving school effectiveness, and the quality of learning and teaching' (World Bank, 2015, pg. 34). As part of the evaluation, half of the 64 treatment schools received this amount ('normal') whereas the other half received twice the amount, 500,000 Nigerian Naira ('high'). The SBMC training contained leadership and school management skills as well as the importance of community involvement when taking decisions. The SBMC was to draw up a School Improvement Plan

before receiving the grant to determine the priority areas in which improvements were deemed necessary for the school.

## Experimental design and timeline

The 128 schools in the study sample were randomly assigned to either the treatment (receiving the grant and SBMC training) or control group (no intervention) with equal probability and the treatment schools were then again randomly divided into the normal grant and the high grant groups with equal probability. After a pilot study that tested the difficulty and wording of the student tests, the collection of baseline data took place in July and August 2018. Subsequently, the intervention was implemented and 14 months later the endline data were collected in November and December 2019. Questions in the endline surveys were adjusted so that the correct school year (2018/19) was referred to. The study did not provide any monetary incentives for participation but rewarded students with a cookie for completing the survey.

## Data

Baseline and endline data were collected via standardised questionnaires by enumerators fluent in the local dialect of Hausa. Answers were recorded by enumerators on tablets. Questions were drawn up in English, translated and back-translated to Hausa and available in Hausa on the tablets. Enumerators were recruited from local communities by the survey firm to ensure familiarity with the local dialect of the Hausa language. They were trained to create an encouraging, trusting and private environment for primary school children and to emphasise that their answers were confidential and would not impact their grades in school. Student interviews were conducted with both the enumerator and the child seated on a mat to make the child feel comfortable. At times, it proved difficult to create a private setting for the surveys, especially for the student tests, because interviews often had to be conducted outside in the shade of school buildings or trees where other curious children could easily pass or watch. Frequent reminders by the supervising team to the enumerators were given to ensure privacy with mixed success.

At each school, five different surveys were administered. First, the team leader of the enumerator group would fill out an observational questionnaire that collected impressions of the infrastructure, people present upon arrival and their activities as well as information copied from the school registries (enrolment and attendance records of students and teachers) if available. Furthermore, three separate questionnaires were designed and administered with each primary school's headmaster, additional teachers and members of their School-Based Management Committees (SBMC). Data on their demographics, attitudes, their perception of the school's challenges, any interventions or trainings the school might have received in the past year, and activities and characteristics of the SBMC were collected.

The pupil surveys started with the numeracy and the literacy test and subsequently collected some information on their demographics and opinions on their primary school. The mathematics and Hausa questions started out with the easiest tasks – counting and recognising single digits or letters – and became progressively more difficult. If a student answered a question incorrectly, a second question of the same difficulty level had to be answered correctly before moving on to the next difficulty level. Failing that, the questionnaire automatically moved on to the next sub-section to avoid frustrating the children with too many questions they were unable to answer.

At baseline, we interviewed 5,717 Grade 2 and 3 students, 88 headmasters, 181 teachers and 285 SBMC members while at endline, we interviewed 6,013 Grade 2, 3 and 4 students, 99 headmasters, 175 teachers and 348 SBMC members. It was not feasible to track and match students from the baseline to the endline survey, so we consider our data as independently pooled cross sections.

## Estimation strategies

The random allocation of schools into treatment and control group allows us to establish a credible counterfactual so that any treatment effects we identify can be causally associated with the intervention. This assumes that randomization was successful at creating control and

treatment groups that are balanced – we display the balance of some observables in Table 2 below.

First, we estimated the average effect of belonging to a school in the treatment group, the intent- to-treat effect (ITT), on each outcome variable Y, using endline data. In the cases where outcome variables were collected only once per school (observations, headmaster survey), the estimation was as follows:

$$Y_i = \alpha_0 + \alpha_1 T_i + \varepsilon_i \quad (1)$$

where $T_i$ is a dummy variable for belonging to a treatment group school $i$ and $\varepsilon_i$ is an error term for school $i$. In the case of the outcome variable having multiple observations per school (students, teachers, SBMC members), standard errors were clustered at the unit of randomization, the school:

$$Y_j = \alpha_0 + \alpha_1 T_i + \varepsilon_{ij} \quad (2)$$

where $j$ refers to an individual surveyed.

In the Appendix, we present two alternative specifications: First, we add the baseline values of the outcome variable as a control variable to our intent-to-treat OLS estimation (ANCOVA). In the case of students' learning achievements where we have several observations per school, standard errors were again clustered at the unit of randomization, the school. In other words, $\varepsilon_i$ is replaced with $\varepsilon_{ij}$:

$$Y_i = \alpha_0 + \alpha_1 T_i + Y_{i(t-1)} + \varepsilon_i \quad (3)$$

where $Y_{i(t-1)}$ is the lagged outcome variable at baseline.

Second, we use a difference-in-differences estimator for which we assume (i) that treatment and control groups were on parallel trends before the introduction of the intervention and (ii) that participants were unable to select into treatment group schools. Both assumptions should be met due to the randomized design of the study.

$$Y_{it} = \alpha_0 + \alpha_1 T_i + \alpha_2 E_t + \alpha_3 T_i * E_t + \varepsilon_{it} \quad (4)$$

where $E_t$is an endline dummy variable equal to 1 if the data point was collected as part of the endline survey, i.e. after the intervention was implemented.

Given that many headmasters and SBMC members at treatment schools did not report receiving a grant at endline, we also run an Instrumental Variable analysis estimating the Complier Average Causal Effect (CACE). The first-stage regressions are:

$$X_i = \beta_0 + \beta_1 T_i + \epsilon_i \quad (5)$$

where $X_i$is either a dummy variable for if the headmaster or any SBMC member mentioned the school receiving any grant or a dummy variable for when anyone mentioned a training and $\epsilon_i$ is an error term for school $i$. As before, we cluster standard errors at the school level for all individual-level outcomes, such as student test scores.

Since we test 13 different education outcomes (see Table 4), there is a heightened probability of falsely rejecting at least one null hypothesis (Anderson, 2008). Hence, we correct standard errors for multiple hypothesis testing using the Benjamini-Hochberg method and report a second set of statistical significance levels in Tables 2, 4, A1 and A3 where we present our main results (Benjamini and Hochberg, 1995).

## 2.4. Results

### Summary statistics and balance

In Table 2, we document our baseline results and the balance checks for whether randomization was successful. The three school types included in the sample – regular, islamiyya and nomadic primary schools – were split quite similarly across treatment and control group. 98 of the 99 headmasters interviewed were Muslim and 94 were male, with 4 female headmasters in the treatment and one in the control group. On average, schools in the study sample had 190 enrolled students, 39% of whom were female. Students' numeracy

and literacy skills at baseline were very poor: Out of a maximum score of 20, Grade 2 students scored 6.2 points in mathematics and 1.6 points in Hausa, while Grade 3 students scored 8.8 points in mathematics and 3.8 points in Hausa. Upon enumerators' unannounced arrival on the day of the baseline survey, on average less than one teacher (0.71) was present and, according to school registries (only available at 108 of the 128 schools), 36% of pupils were absent. The observed teacher absenteeism is substantially worse than the 17% of teachers absent from school observed by Bold et al. (2017) in Nigeria, though they surveyed primary schools in different states (Anambra, Bauchi, Ekiti and Niger) with at least one grade 4 class.

The infrastructure and equipment of the school was also recorded and summarized in three indices ranging from 0 to 1 (shown in the first three lines of Table 2 below). Some examples of the more detailed measurements of the schools' learning environments, sanitation and facilities used in creating the indices are presented in Table 3. To create the indices, all measures were scaled to the same range and a simple average was calculated. Table 3 shows a list of all variables included in the indices.

Schools had on average 3 classrooms but for example working material and school uniforms were unavailable at the large majority of sample schools. Only 11% of schools had any kind of water supply, only 22% had any toilets and only 5% had access to electricity.

The randomization produced a treatment and control group that were on average balanced across outcome measures with three exceptions. Treatment schools were more likely to have any toilets available (p-value 0.08), less likely to have any students engaged in any learning activities[11] (p-value 0.04), and their Grade 2 students performed worse on the literacy test (p-value 0.06).

Even though these significance levels do not hold up to multiple hypothesis testing, we choose to run specifications different to our standard ITT OLS specification in these three cases. For the first two unbalanced outcome variables (toilets, and any learning taking place), we add the schools' baseline values of the two outcomes as control variables in our regressions that

[11] The dummy variable 'any learning taking place' measures teacher absenteeism from the classroom where the value 1 means that enumerators observed students engaged in some learning activity upon their unannounced arrival.

we then report in Table 4. This ANCOVA estimation strategy was outlined in section 2.4 and is implemented as a robustness check for all the balanced outcome variables as well in the

*Table 2: Sample characteristics and balance checks*

| | Sample | Control | | Treatment | | Difference in means (C-T) | p-value of difference in means |
|---|---|---|---|---|---|---|---|
| | | N (1) | Mean (2) | N (3) | Mean (4) | (5) | (6) |
| **School type** | | | | | | | |
| Regular | 106 | 52 | | 54 | | -2 | 0.64 |
| Islamiyya | 16 | 10 | | 6 | | 4 | 0.29 |
| Nomadic | 6 | 2 | | 4 | | -2 | 0.41 |
| **Summary indices** | | | | | | | |
| Quality of the learning environment[1] | 0.336 | 64 | 0.343 | 64 | 0.323 | 0.021 | 0.45 |
| Sanitation[1] | 0.219 | 64 | 0.240 | 64 | 0.188 | 0.052 | 0.27 |
| School facilities[1] | 0.152 | 64 | 0.157 | 64 | 0.150 | 0.007 | 0.82 |
| Any toilets | 0.220 | 63 | 0.286 | 64 | 0.156 | 0.129 | 0.08* |
| **Pupil enrolment** | | | | | | | |
| Total enrolment | 190 | 64 | 166 | 64 | 214 | -48 | 0.27 |
| Female enrolment rate | 0.394 | 54 | 0.397 | 54 | 0.391 | 0.006 | 0.82 |
| **Pupil attendance** | | | | | | | |
| No. students observed in Grade 2 / 3 | 53 | 64 | 55 | 64 | 50 | 4.4 | 0.60 |
| Pupil absence rate according to registry | 0.356 | 54 | 0.310 | 54 | 0.403 | -0.093 | 0.15 |
| **Teacher attendance** | | | | | | | |
| Number of teachers present | 0.711 | 64 | 0.859 | 64 | 0.563 | 0.297 | 0.22 |
| **Lessons** | | | | | | | |
| Any learning taking place | 0.234 | 64 | 0.313 | 64 | 0.156 | 0.156 | 0.04** |
| **Learning achievement scores (out of 20)** | | | | | | | |
| Grade 2 math | 6.22 | 1,542 | 6.29 | 1,701 | 6.16 | 0.131 | 0.33 |
| Grade 3 math | 8.79 | 1,299 | 8.79 | 1,163 | 8.78 | 0.010 | 0.96 |
| Grade 2 literacy | 1.61 | 1,542 | 1.73 | 1,701 | 1.51 | 0.227 | 0.06* |
| Grade 2 school means literacy | 1.58 | 59 | 1.71 | 62 | 1.46 | 0.251 | 0.47 |
| Grade 3 literacy | 3.77 | 1,299 | 3.84 | 1,163 | 3.70 | 0.215 | 0.52 |

Notes: [1]These variables are indices ranging from 0 to 1 with 0 representing a poor and 1 a good outcome. Statistical significance levels (10%, 5%, 1%) based on naïve p-values represented with */**/*** and based on Benjamini-Hochberg adjusted p-values corresponding to a 5% significance level represented with ‡.

*Table 3: Input variables selected for three summary indices*

| | Baseline mean (1) |
|---|---|
| **Quality of the learning environment** | |
| Condition of the school building[1] | 2.52 |
| Number of classrooms | 3.04 |
| Condition of classrooms[1] | 2.49 |
| Any blackboard in the classroom | 0.489 |
| Benches, chairs and tables[2] | 0.36 |
| Books[3] | 3.87 |
| Working material[3] | 5.34 |
| Any educational posters | 0.052 |
| Pupil uniforms[3] | 4.92 |
| | |
| **Sanitation** | |
| Any water supply | 0.11 |
| Any toilets | 0.22 |
| Any faeces around the compound | 0.31 |
| | |
| **School facilities** | |
| Any headmaster's office | 0.50 |
| Any staff room | 0.14 |
| Any storage room for learning materials | 0.09 |
| Any power supply | 0.05 |
| | |
| Number of observations | 128 |

Notes: [1] Scale of 1 to 5 where 1 is very poor and 5 very good.
[2] Scale of 1 to 5 where 1 is sufficiently available for all pupils, 2 is 'All pupils seated but more children per chair/bench than designated, 3 is 'More than half of pupils sit on chairs/benches', 4 is 'Less than half …' and 5 is 'No chairs / benches available'.
[3] Scale of 1 to 6 where 1 is equivalent to 'more than ¾', 2 corresponds to '¾ to 1/2', 3 corresponds to 'One half', 4 corresponds to '¼ to ½', 5 corresponds to 'less than ¼' and 6 corresponds to 'None'.

appendix. For the third unbalanced outcome, Grade 2 pupils' literacy, we aggregate pupils' test scores at the school level and use these average school-level literacy test scores, which is balanced at baseline (p-value 0.47).

## Impacts on infrastructure, educational attainment and learning outcomes

Results for the outcomes on infrastructure, enrolment, attendance and learning outcomes are reported in Table 4. For each outcome we present the intent-to-treat estimates in column (1) and the instrumental variable estimates in columns (2) and (3). With one exception, the intervention had no discernible impact across outcomes and specifications. This proves robust to the alternative ANCOVA and difference-in-differences (DiD) specifications (Appendix Table

*Table 4: ITT and CACE estimates for outcomes of interest*

| | ITT OLS (1) | CACE IV (grant) (2) | CACE IV (training) (3) |
|---|---|---|---|
| **First stage for IV estimators** | | | |
| Treatment assignment | | 0.672*** | 0.377*** |
| | | (0.062) | (0.064) |
| *F-statistic* | | 118.93 | 35.12 |
| | | | |
| **Infrastructure and equipment** | | | |
| Quality of the learning | 0.036 | 0.061 | 0.109 |
| environment[1] | (0.031) | (0.046) | (0.082) |
| Sanitation[1] | 0.069 | 0.109 | 0.199 |
| | (0.048) | (0.074) | (0.134) |
| School facilities[1] | -0.009 | -0.023 | -0.042 |
| | (0.035) | (0.055) | (0.100) |
| Any toilets[2] | 0.232***/‡ | 0.212* | 0.386* |
| | (0.073) | (0.125) | (0.228) |
| | | | |
| **Pupil enrolment** | | | |
| Total enrolment | 51.96 | 75.2 | 134 |
| | (41.65) | (63.0) | (112) |
| Female enrolment rate | 0.0094 | 0.017 | 0.028 |
| | (0.041) | (0.054) | (0.090) |
| | | | |
| **Pupil attendance** | | | |
| Any students present | -0.059 | -0.091 | -0.249 |
| | (0.084) | (0.085) | (0.236) |
| | | | |
| **Teacher attendance** | | | |
| Any teacher present | -0.073 | -0.155 | -0.282 |
| | (0.090) | (0.137) | (0.253) |
| Any learning taking place[2] | 0.020 | -0.014 | -0.025 |
| | (0.081) | (0.123) | (0.223) |
| | | | |
| **Normalised learning achievement scores** | | | |
| Grade 2 numeracy | -0.069 | -0.125 | -0.233 |
| | (0.100) | (0.166) | (0.314) |
| Grade 3 numeracy | -0.036 | -0.140 | -0.234 |
| | (0.126) | (0.199) | (0.334) |
| Grade 2 school means literacy | -0.067 | -0.137 | -0.238 |
| | (0.104) | (0.166) | (0.295) |
| Grade 3 literacy | -0.159 | -0.316 | -0.528 |
| | (0.131) | (0.213) | (0.373) |

Notes: [1]These variables are indices ranging from 0 to 1 with 0 representing a poor and 1 a good outcome.
[2]These variables were not balanced at baseline; the reported ITT OLS result is from running an ANCOVA specification.
Standard errors in parentheses. Statistical significance levels (10%, 5%, 1%) based on naïve p-values represented with */**/*** and based on Benjamini-Hochberg adjusted p-values corresponding to a 5% significance level represented with ‡. The first-stage results of the instrumental variable specification in columns (2) and (3) were excluded from multiple hypothesis testing corrections.
Sample sizes: School-level regressions (first stage, infrastructure and equipment, pupil enrolment, both attendance measures, and Grade 2 literacy) had between 115 and 125 observations. Student-level regressions (normalized learning achievement scores) had between 1901 and 2770 observations. Exception: Female enrolment rate had 64 or 65 observations. Also note columns (1) and (3) of Table 2.

A1) and to differentiating between the normal and high grant amount treatment schools (Appendix Table A3).

The mentioned exception is the dummy variable relating to whether schools had any toilets. Table 4 shows that the ANCOVA and CACE estimates are significant at the conventional levels, as are the DiD estimates (Table A1). However, this statistical significance disappears when correcting for multiple hypothesis testing with the exception of the ANCOVA specification where the positive impact of the intervention on school's toilets proves robust to the Benjamini-Hochberg correction. We therefore treat this treatment effect as anecdotal evidence.

## 2.5. Discussion of potential mechanisms

To understand why the intervention had no discernible impact on primary schools' learning environment or achievement, we explore five different channels to explain the null results. The auxiliary analyses are presented below.

### Implementation challenges

First, the implementation of the grants and SBMC trainings may have been faulty from the responsible NIPEP office in Sokoto so that many treatment schools may have never received any money or trainings. This is what Reinikka and Svensson (2011) find in Uganda where less access to public information about school grants (i.e. newspapers) led to more capture of school grants by local governments.

We use the number of respondents that reported whether any intervention took place, a grant was disbursed or a training was offered as an indication for potential challenges in the implementation and find that the reporting of the intervention was very low.

At endline, at only 32 of the 64 treatment schools did any SBMC member report receiving an intervention, and at only 27 treatment schools did any SBMC member report receiving a

grant. Similarly, only 26 of the 52 headmasters interviewed at treatment schools reported an intervention and only 17 reported receiving a grant. Many respondents also did not report any training for the SBMC committee (Table 5). The first stage of the instrumental variable estimations (CACE), reported in Table 4 above, showed that at 66.7% of treatment schools at least one respondent reported an intervention and at 36.7% at least one respondent reported a training. However, the CACE estimations echo the ITT estimations regarding the outcomes of interest so that we conclude that implementation challenges only play a minor role in explaining the null results.

## Spending challenges

Schools may have spent the grants on school-unrelated matters which could be another reason why the intervention did not improve learning environments or students' learning achievements. This could have happened as a consequence of corrupt practices by those SBMC members with access to the school bank account, but it could also have happened with the best intentions as a consequence of insufficient communication with and training of SBMC members on the intended outcomes of the project (possibly there was an implementation lag until SBMC members learned about the benefits of school improvements and the endline happened too early to determine positive impact, see King and Behrman 2009).

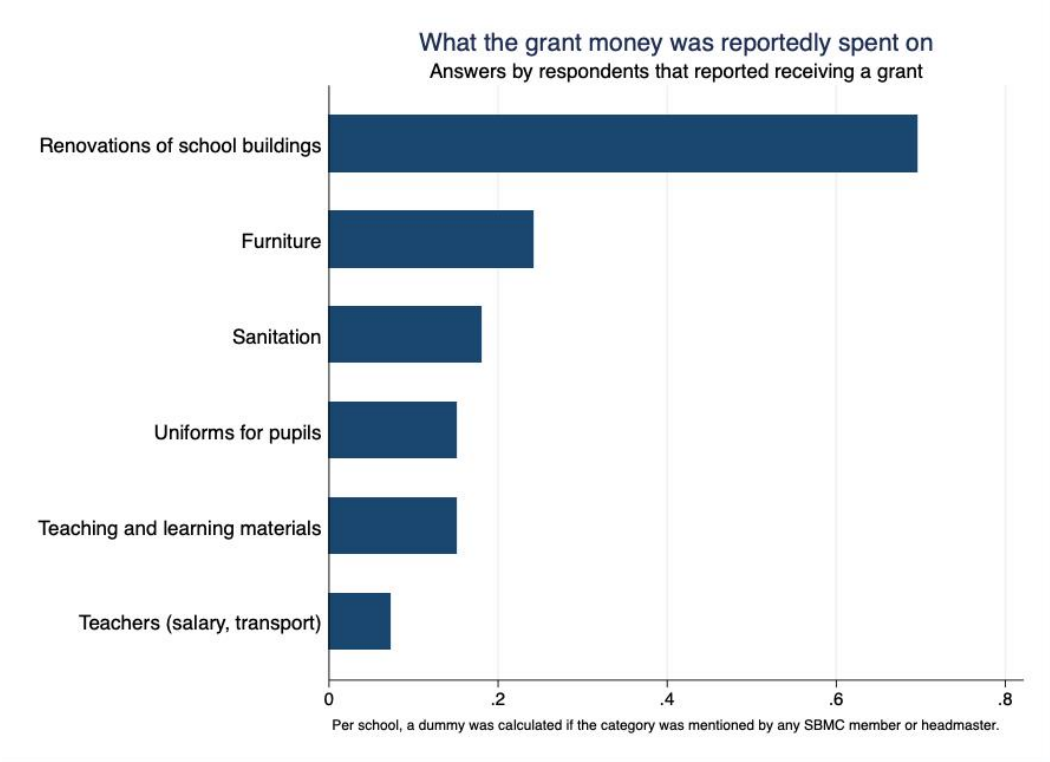
To verify this explanation, we look at what SBMC members and headmasters claim to have spent the grant money on at the treatment schools where at least one person reported receiving a grant (displayed in Figure 1 below) and verify their reported usage of the grant by testing for changes in the reported spending domain from base- to endline survey.

The most frequently cited use of the grant money was renovations (70%) and furniture (24%). These reported grant usages need to be interpreted with caution, though, as only 65 headmasters or SBMC members at treatment schools reported a grant in the endline survey, creating a non-random subsample of only 33 out of the 64 treatment schools.

*Table 5: Reporting of intervention components by headmasters and SBMC members*

| **Subsample: Treatment schools at endline** | Likelihood of reporting an **intervention** (1) | Likelihood of reporting a **grant** (2) | Likelihood of reporting a **SBMC training** (3) | Number of respondents (4) |
|---|---|---|---|---|
| Headmaster | 50.0% | 32.7% | 38.5% | 52 |
| SBMC Chairman | 50.0% | 36.0% | 26.0% | 50 |
| SBMC Vice Chair | 28.6% | 35.7% | 35.7% | 14 |
| SBMC Secretary | 61.5% | 53.8% | 15.4% | 13 |
| SBMC Treasurer | 50.0% | 50.0% | 12.5% | 8 |
| SBMC Woman Leader | 40.0% | 27.3% | 18.2% | 11 |
| SBMC Pupil representative | 0.0% | 0.0% | 100.0% | 1 |
| SBMC Ordinary member | 46.6% | 32.8% | 17.2% | 64 |

*Figure 1: Frequency of grant use mentions (at treatment schools where grant was reported)*




To verify whether spending the grant money in the reported way was visible in our data, we run two sets of regressions. As our sample is non-random, we cannot use an ITT OLS specification, as the treatment schools that report using the grant for renovations are likely to be systematically different to the treatment schools that do not, creating bias. Therefore, our first verification uses an OLS regression which includes the baseline value of the

independent variable as a control variable (ANCOVA). Second, we use a difference-in-differences estimator:

Column (1) of Table 6 reports $\alpha_1$: $Y_i = \alpha_0 + \alpha_1 X_i + Y_{i(t-1)} + \varepsilon_i$
Column (2) of Table 6 reports $\beta_3$: $Y_{it} = \beta_0 + \beta_1 X_i + \beta_2 E_t + \beta_3 X_i * E_t + \varepsilon_{it}$

where $Y_{i(t-1)}$ is the lagged outcome variable at baseline, $X_i$ is the independent variable listed in Table 6's header cells and $E_t$ is an endline dummy variable equal to 1 if the data point was collected as part of the endline survey, ie after the intervention was implemented.

The small number of schools left in this subsample make these regression results merely suggestive as they are based on few observations, eg n = 23 when the independent variable is 'reported using the grant for renovations'.

Table 6 does not give us confidence in the truthfulness of the headmasters and SBMC members' responses. Whenever respondents claimed for the school to have received a grant, their reported use of the grant money does not translate into improvements in the specified domain. For instance, if schools reported spending the grant on renovations, we do not observe any improvement in the condition of the school building, classrooms, doors and windows from base- to endline survey. One interesting exception is that schools that claimed to have used the grant to build toilets are in fact significantly more likely to have any toilet facility at the time of the endline survey.

*Table 6: Comparison of reported grant uses with enumerator observations*

| **Subsample:**<br>**Treatment schools where at least one respondent reported a grant** | ANCOVA<br>(1) | DiD<br>(2) |
|---|---|---|
| **Independent variable: Reported using grant for renovations (dummy)** | | |
| General condition of school buildings / compound[1] | -0.124 | 0.079 |
| | (0.103) | (0.129) |
| Condition of classrooms[1] | -0.073 | 0.050 |
| | (0.091) | (0.120) |
| Condition of doors[1] | -0.090 | 0.053 |
| | (0.100) | (0.142) |
| Condition of windows[1] | -0.022 | 0.081 |
| | (0.103) | (0.145) |
| **Independent variable: Reported using grant for learning materials (dummy)** | | |
| Availability of learning materials[1] | 0.358 | 0.453 |
| | (0.211) | (0.334) |
| **Independent variable: Reported using grant for sanitation (dummy)** | | |
| Sanitation[1] | 0.123 | 0.213 |
| | (0.135) | (0.191) |
| Any toilet (dummy) | 0.603*** | 0.631** |
| | (0.170) | (0.289) |
| **Independent variable: Reported using grant for uniforms (dummy)** | | |
| Pupil uniforms[1] | -0.198 | -0.375 |
| | (0.405) | (0.368) |

Notes: [1] These are indices ranging from 0 (very poor) to 1 (very good).
Standard errors in parentheses. Statistical significance levels (10%, 5%, 1%) based on naïve p-values represented with */**/***.
Sample size: 32 schools. Exception: Availability of learning materials and uniforms: 6 schools

## Insufficient grant amount

If the amount of the grant was insufficient to eliminate the constraints to delivering a quality education, this could explain the zero impact of the intervention as well. Given the very lacking and, if existent, deficient infrastructure of our sample schools (see Table 3 in section 3.1.), the grant of approximately US$ 2,250 could not be enough to make a lasting improvement in the infrastructure such that educational attainment and learning outcomes are affected. The treatment schools that received twice the grant amount, however, also recorded no improvements along infrastructural or educational dimensions (see Table A3 in the Appendix). This would suggest that even US$ 4,500 were not enough to significantly improve the deficient delivery of education.

## Teacher absenteeism

Instead of schools' financial situation, the binding constraint in delivering quality primary education in the study context could also lie elsewhere: teacher absenteeism. Upon enumerators' arrival for the endline survey, there was no learning taking place at 74% of schools and at 45%, no teacher was even present. In other words, learning can take place anywhere, even in the shade of the tree, but without a teacher present and giving lessons, students are unlikely to learn. Disbursing grants to schools where teachers are regularly missing would therefore not translate into improvements in learning outcomes, either. However, this would not explain why we do not see any improvements along the other main outcome variables, such as infrastructure or learning materials, as the SBMC could have invested in these without teachers being present.

## Capacity of SBMC

Following Blimpo, Evans and Lahire's (2015) argument of the importance of SBMC's local capacity, we assess whether the average literacy[12] of the SBMC members interviewed limited the potential impact of the intervention by interacting it with a treatment dummy.

$$Y_i = \gamma_0 + \gamma_1 T_i + \gamma_2 L_i + \gamma_3 T_i * L_i + \varepsilon_i \quad (6)$$

where $T_i$ is a dummy variable for belonging to a treatment group school $i$, $L_i$ is the average literacy rate of interviewed SBMC members interviewed at school $i$, and $\varepsilon_i$ is an error term. As before, we cluster standard errors at the school level for all outcomes that were measured at the school level.[13]

Table 7 below displays the estimates of the coefficients of interest, $\gamma_3$. We find no heterogeneous effects of the intervention by SBMC capacity. In fact, both pupil and teacher

[12] This measure was constructed as the percent of interviewed SBMC members that confirmed being able to read a letter.

[13] Using an alternative specification (not shown) where we include the baseline values of the outcome variable as a control variable (ANCOVA) yields the same null results, with the same two exceptions of negative correlations as in the endline OLS specification reported in Table 7.

attendance even show up as being negatively impacted by higher SBMC literacy rates and receiving the NIPEP intervention. Blimpo, Evans and Lahire (2015) estimated that a minimum of 45% adult literacy was needed for their intervention to show effects on students' learning outcomes. As only 44% of SBMC members interviewed at endline said they were able to read a letter, SBMC capacity in our study context may indeed fall below that critical literacy threshold. Therefore, we do not disprove Blimpo, Evans and Lahire's (2015) finding but are

*Table 7: Interacting SBMC literacy with treatment assignment*

| | (1) |
|---|---|
| **Infrastructure and equipment** | |
| Quality of the learning environment[1] | -0.049 |
| | (0.083) |
| Sanitation[1] | 0.008 |
| | (0.133) |
| School facilities[1] | -0.022 |
| | (0.090) |
| **Pupil enrolment** | |
| Total enrolment | 72.69 |
| | (113.6) |
| Female enrolment rate | -0.079 |
| | (0.113) |
| **Pupil attendance** | |
| Any students present | -0.385* |
| | (0.223) |
| **Teacher attendance** | |
| Number of teachers present | 0.174 |
| | (1.56) |
| Learning taking place | -0.519** |
| | (0.211) |
| **Normalised learning achievement scores** | |
| Grade 2 numeracy | 0.091 |
| | (0.284) |
| Grade 3 numeracy | -0.345 |
| | (0.330) |
| Grade 2 school means literacy | 0.249 |
| | (0.288) |
| Grade 3 literacy | -0.003 |
| | (0.371) |

Notes: Standard errors in parentheses. Statistical significance levels (10%, 5%, 1%) based on naïve p-values represented with */**/***.

Sample sizes: School-level regressions (infrastructure and equipment, total enrolment, pupil attendance, learning taking place and Grade 2 literacy) had 114 observations. Student-level regressions (normalized learning achievement scores) had between 1853 and 2501 observations. Exceptions: 61 schools for female enrolment rate and 30 observations for the number of teachers regressions.

simultaneously unable to confirm the importance of local capacity for the delivery of education in our study context.

Pradhan and De Ree (2014) provide an alternative channel through which SBMC capacity could be limiting the impact of training plus school grant interventions. Indonesian primary school pupils had substantial learning gains when school committees were democratically elected or had joint planning meetings with the village council. In the Nigerian context, this might mean that learning outcomes could have potentially been improved by implementing democratic SBMC elections and facilitating joint meetings with the village elders.

## 2.6.Conclusion

We examined the impact of the grant and school-based management components of the Nigerian Partnership of Education Project (NIPEP) on primary school's infrastructure, educational attainment and learning achievement in Sokoto state, Nigeria. NIPEP was a USD 100 million project funded by the World Bank and the Global Partnership for Education from 2015 to 2019 in five states in Northern Nigeria with the goal of improving access to quality primary school education with a special focus on girls. The grant disbursal component alone cost approximately USD 21.1 million.

Other school-based management interventions in combination with grant disbursals usually have not improved educational attainment or learning outcomes (Banerjee et al. 2010; Glewwe and Maïga 2011; Blimpo, Evans, and Lahire 2015; Lassibille et al. 2010; Pradhan and De Ree 2014; Santibañez, Abreu-Lastra, and O'Donoghue 2014), though only few studies involved large-scale interventions like ours: Garcia-Moreno, Gertler, and Patrinos (2019) and Santibañez, Abreu-Lastra, and O'Donoghue (2014) evaluate large-scale school-based management interventions plus grant disbursals in Mexico, while the remaining studies evaluated small pilot projects or their own experiment. Here, we present novel insights from a large-scale intervention in a lower-middle income and sub-Saharan African context, Nigeria, that proved entirely ineffective at delivering impact. One year after implementation, the

intervention appears to have had no impact on schools' infrastructure, educational attainment or learning outcomes.

There are several possible explanations for this lack of impact. First, the implementation of the intervention may have been faulty from the responsible state's education authorities, which is corroborated by the low level of reporting an intervention by respondents at treatment schools. Second, schools may have spent the grant on school-unrelated matters while reporting untrue uses of the grant at the endline survey. Anecdotal evidence from a subsample of treatment schools shows that no improvements could be detected in the domains that respondents claimed to have invested the grant in. Third, the grant amount of 250,000 Nigerian Naira (approx. PPP-adjusted int-$ 2,250) may have been insufficient to alleviate the infrastructural and working material deficits of the primary schools in the sample. Fourth, the binding constraint in educational attainment and learning achievement could be teacher absenteeism, rather than a lack of money. Upon arrival for the endline survey, there were no teachers present at 45% of sample schools and no learning taking place at 74%. Fifth, the local capacity of SBMC members may have been limiting the committees' ability to transform the training's lessons and grant into measurable school improvements. Our proxy measure of SBMC capacity, members' mean literacy level, proved quite low (44% of interviewed SBMC members at endline reported being able to read a letter) but was also uncorrelated with the success of the intervention at treatment schools.

Despite its poor record at generating measurable, meaningful and sustainable improvements in education outcomes, school-based management interventions in combination with grant disbursals to schools have recently gained popularity with funders and policy-makers. Our evidence underlines the importance of further research to understand the determinants of the positive outlier SBM interventions' success. For example, identifying the complementary input factors, such as sufficient ministerial and SBMC capacity, e.g. its connections to the village council (Pradhan and De Ree 2014), or satisfactory teacher attendance and motivation, and context-specific characteristics that make SBM and grant disbursal interventions with positive impact possible is important to understand for researchers as well as policy-makers.

# ESSAY TWO

## 3. Diagnostic testing for hypertension, diabetes and hypercholesterolaemia in low-income and middle-income countries: a cross-sectional study of data from 994 185 individuals from 57 nationally representative surveys

Joint work with:

Isabelle von Polenz, Maja-Emilia Marcus, Michaela Theilmann, David Flood, Kokou Agoudavi, Krishna Kumar Aryal, Silver Bahendeka, Brice Bicaba, Pascal Bovet, Luisa Campos Caldeira Brant, Deborah Carvalho Malta, Albertino Damasceno, Farshad Farzadfar, Gladwell Gathecha, Ali Ghanbari, Mongal Gurung, David Guwatudde, Corine Houehanou, Dismand Houinato, Nahla Hwalla, Jutta Adelin Jorgensen, Khem B Karki, Nuno Lunet, Joao Martins, Mary Mayige, Sahar Saeedi Moghaddam, Omar Mwalim, Kibachio Joseph Mwangi, Bolormaa Norov, Sarah Quesnel-Crooks, Negar Rezaei, Abla M Sibai, Lela Sturua, Lindiwe Tsabedze, Roy Wong-McClure, Justine Davies, Pascal Geldsetzer, Till Bärnighausen, Rifat Atun, Jennifer Manne-Goehler, Sebastian Vollmer

Published in the Lancet Global Health:

Ochmann, S., von Polenz, I., Marcus, M.E., Theilmann, M., Flood, D., Agoudavi, K., Aryal, K.K., Bahendeka, S., et al. 2023. Diagnostic testing for hypertension, diabetes, and hypercholesterolaemia in low-income and middle-income countries: a cross-sectional study of data for 994 185 individuals from 57 nationally representative surveys. *The Lancet Global Health*, *11*(9), pp.e1363-e1371.

## 3.1. Abstract

**Background:** Testing for the risk factors of cardiovascular disease, which include hypertension, diabetes, and hypercholesterolaemia, is important for timely and effective risk management. Yet few studies have quantified and analyzed testing of cardiovascular risk factors in low-income and middle-income countries (LMICs) with respect to sociodemographic inequalities. We aimed to address this knowledge gap.

**Methods:** In this cross-sectional analysis, we pooled individual-level data for non-pregnant adults aged 18 years or older from nationally representative surveys done between 1 January 2010 and 31 December 2019 in LMICs that included a question about whether respondents had ever had their blood pressure, glucose, or cholesterol measured. We analyzed diagnostic testing performance by quantifying the overall proportion of people who had ever been tested for these cardiovascular risk factors and the proportion of individuals who met the diagnostic testing criteria in the WHO package of essential noncommunicable disease interventions for primary care (PEN) guidelines (ie a BMI > 30 kg/$m^2$ or a BMI >25 kg/$m^2$ among people aged 40 years or older). We disaggregated and compared diagnostic testing performance by sex, wealth quintile, and education using two-sided *t* tests and multivariable logistic regression models.

**Findings:** Our sample included data for 994 185 people from 57 surveys. 19·1% (95% CI 18·5–19·8) of the 943 259 people in the hypertension sample met the WHO PEN criteria for diagnostic testing, of whom 78·6% (77·8–79·2) were tested. 23·8% (23·4–24·3) of the 225 707 people in the diabetes sample met the WHO PEN criteria for diagnostic testing, of whom 44·9% (43·7–46·2) were tested. Finally, 27·4% (26·3–28·6) of the 250 573 people in the hypercholesterolaemia sample met the WHO PEN criteria for diagnostic testing, of whom 39·7% (37·1–2·4) were tested. Women were more likely than men to get tested for all three risk factors, as were people in higher wealth quintiles compared with those in the lowest wealth quintile and people with at least secondary education compared with those with less than primary education.

**Interpretation:** Our study shows opportunities for health systems in LMICs to improve the targeting of diagnostic testing for cardiovascular risk factors and adherence to diagnostic testing guidelines. Risk-factor-based testing recommendations rather than sociodemographic characteristics should determine which individuals are tested.

## 3.2. Introduction

For the past three decades, cardiovascular disease has been the most common cause of death in low-income and middle-income countries (LMICs) (IHME, 2023). However, diagnostic testing for risk factors for cardiovascular disease in LMICs has remained low, and 56-69% of adults with one of the three major risk factors for cardiovascular disease — ie hypertension, diabetes, and hypercholesterolaemia — are undiagnosed (Geldsetzer et al., 2019; Manne-Goehler et al., 2019; Flood et al., 2021; Marcus et al., 2021).

Studies in LMICs have assessed the performance of health systems in terms of testing for, diagnosing, treating, and controlling cardiovascular risk factors among subpopulations with cardiovascular risk factors (Geldsetzer et al., 2019; Manne-Goehler et al., 2019; Flood et al., 2021; Marcus et al., 2021; NCD Risk Factor Collaboration, 2021), but a study of diagnostic testing for these risk factors in the overall populations of LMICs is lacking. In these previous studies, coverage of testing was the largest gap in the care continuum, and this low coverage majorly inhibited the provision of effective, efficient, and timely management of cardiovascular disease. The *Lancet* Commission on diagnostics explicitly identified a lack of diagnostic capacities as the major driver of low testing coverage in LMICs (Fleming et al., 2021; Wilson et al., 2019). The UN's Sustainable Development Goals similarly provided an impetus to expand universal health coverage in LMICs and address diagnostics gaps (UN, 2023).

Beyond diagnostic testing coverage, analysis of health systems' testing performance should include measures of testing necessity and equity. We define high performance of a health system with regard to diagnostic testing as adherence to the international WHO package of essential noncommunicable disease interventions for primary care (PEN) guidelines, which recommend prioritization of testing of people at high risk of developing cardiovascular disease (WHO, 2020). Robust evidence for the performance of health systems in LMICs in terms of testing for hypertension, diabetes, and hypercholesterolaemia and effective and equitable adherence to the WHO PEN guidelines has not previously been published. Furthermore, no previous studies have used individual-level data that link testing data with cardiovascular risk

and sociodemographic factors to examine access to testing access within health systems in LMICs.

In this study, we analyzed diagnostic testing performance for hypertension, diabetes, and hypercholesterolaemia in the health systems of 56 LMICs. We estimated self-reported diagnostic testing and fulfilment of the WHO PEN testing criteria, assessed whether people who met the WHO PEN criteria for testing were actually tested, and analyzed how diagnostic testing performance differed by sex, wealth, and education.

## 3.3. Methods

### Study design and participants

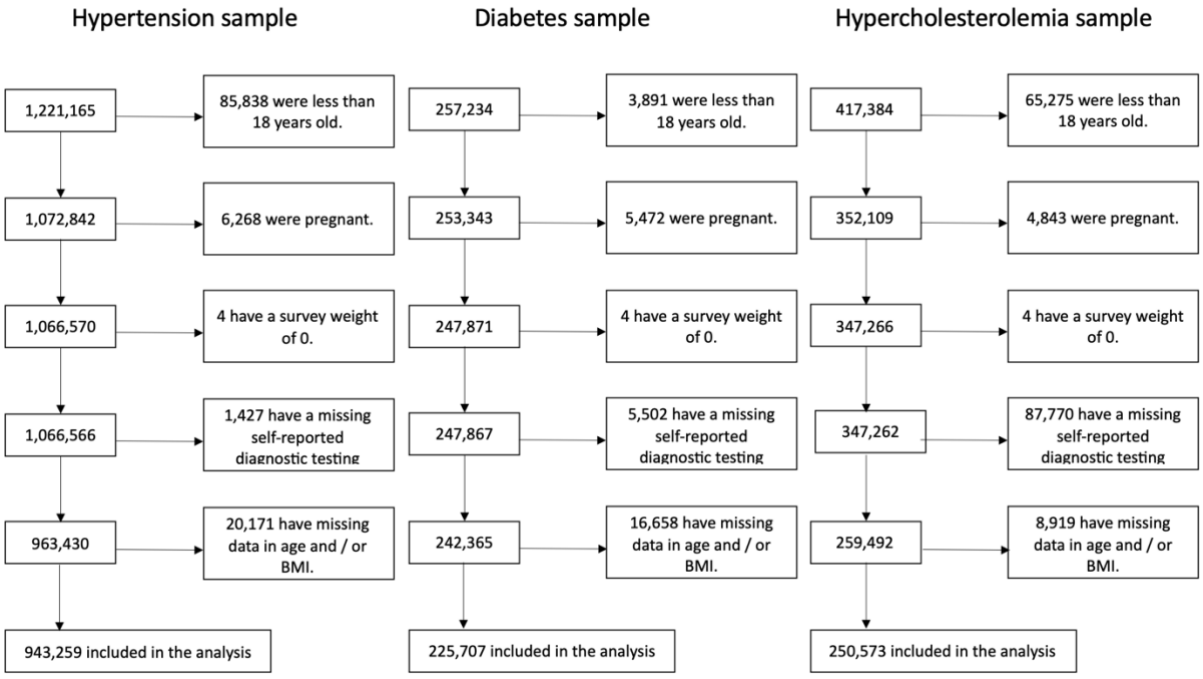
In this cross-sectional analysis, we used pooled, individual-level data from nationally representative surveys in LMICs. Our methods for identifying and pooling surveys followed the same procedure used in previous studies (Geldsetzer et al., 2019; Manne-Goehler et al., 2019; Flood et al., 2021; Marcus et al., 2021; Manne-Goehler et al., 2022) and comprised three parts. First, WHO STEPwise approach to surveillance surveys were identified. Second, we searched six survey resources: the Demographic and Health Surveys (USAID, 2023), the WHO Study on Global Ageing and Adult Health (WHO, 2023), the Gateway to Global Aging studies (University of Southern California Dornsife Center for Economic and Social Research, 2023), the Non-Communicable Disease Risk-Factor Collaboration (NCD-RisC, 2023), the Global Health Data Exchange (IHME, 2023), and the International Diabetes Federation Diabetes Atlas (International Diabetes Federation, 2023). Finally, we did a systematic Google search in April, 2020. Details of the survey inclusion process and the parameters of our Google search are described in appendices 2.1 and 2.2.

To be included, surveys had to be nationally representative, provide individual-level data, have been done between Jan 1, 2010, and Dec 31, 2019 in an LMIC (as classified by the World Bank (2023b) in the survey year), and have contained a question about whether respondents had ever had their blood pressure, glucose, or cholesterol measured. If we were able to access more than one survey for an LMIC, we used the most recent survey.

We included all individuals who were aged 18 years or older, were not pregnant at the time of the survey, had a non-zero survey weight (as determined by the survey team), and had available data for age and BMI. Pregnancy was an exclusion criterion because of the increased probability of undergoing cardiovascular risk factor testing as part of antenatal screening. We created one analysis sample for each risk factor — ie hypertension, diabetes, and hypercholesterolaemia — because not all surveys included questions about whether respondents had ever had their blood pressure, blood glucose, and cholesterol measured, and the varying numbers of respondents with missing information for self-reported diagnostic testing status (figure 3).

Because our analysis included only previously published data, ethics approval was not required. Each included survey received ethical clearance from the relevant country's ethics review committee before data collection, and all participants consented to the use of their data. The Global Health and Population Project on Access to Care for Cardiometabolic Diseases dataset used in this study was deidentified and therefore deemed Non-Human Subjects Research by the institutional review board of the Harvard T H Chan School of Public Health in 2018 (#IRB16-1915). As such, it was exempted from the need for additional ethics approval.

*Figure 2: Selection of hypertension, diabetes, and hypercholesterolaemia sample*



## Outcomes and definitions

We based our analyses on two key outcomes: respondents' diagnostic testing status and fulfilling of the diagnostic testing criteria recommended by the WHO PEN guidelines. For testing status, we constructed three dummy variables for respondents' answers (yes vs no) to the question of whether they had had their blood pressure, blood glucose, or cholesterol (ie, fat in the blood) measured by a doctor or health worker. We used the WHO PEN guidelines (WHO, 2020) because they are the international standard for diagnostic testing of non-communicable diseases and to enable cross-country comparisons. More specifically, because there were no global diagnostic testing criteria for all three risk factors and no separate guidelines specifically for hypertension or hypercholesterolaemia, we used the guidelines in the WHO PEN chapter on diabetes (WHO, 2020) as the universal testing benchmark for all three risk factors. The WHO PEN diabetes guidelines recommend testing people who show symptoms of diabetes, who have a BMI greater than 30 $kg/m^2$, or who are aged 40 years or older and have a BMI greater than 25 $kg/m^2$ (WHO, 2020). However, the included surveys did not collect information on respondents' symptoms (at the time of testing), so we could not consider this criterion in our subsequent analyses. WHO PEN includes guidelines specifically for cardiovascular disease (WHO, 2020), but we did not use them because the recommendations relied on already knowing individuals' hypertension and diabetes status and did not include guidance about who to test for all three risk factors analyzed in this study.

Respondents' sex, height, and weight were recorded by survey administrators. We used height and weight measurements to calculate BMI. We included data for age, education status, and household wealth to study the relationship between these variables and diagnostic testing performance. We grouped education into three categories (less than primary education, at least primary but less than secondary education, and secondary education or more), and used household income or asset ownership data to calculate household wealth quintiles within each country (appendix 2.6). Nine surveys did not include data for household assets or income (appendix 2.4) and were therefore excluded from the analysis of wealth inequalities in access to diagnostic testing.

## Statistical analysis

In each analytic sample, we first assessed the extent to which meeting the diagnostic testing recommendations in the WHO PEN guidelines and testing status overlapped. Second, we ran these analyses disaggregated by sociodemographic characteristics (ie, sex, wealth quintiles, and education status) and tested the significance of the differences between sociodemographic subgroups with a two-tailed, independent, two-sample *t* test for unpaired data with unequal variances. We reported p values after correcting for multiple hypothesis testing according to Benjamini and Hochberg (1995). We additionally calculated 95% CIs for all percentages reported. To corroborate the robustness of the association between sociodemographic characteristics and testing performance, we ran three multivariable logistic regression models, one for each risk factor, among the subsets of people who met the WHO PEN criteria, and reported the findings as odds ratios. Testing status served as the binary outcome variable, and sex, wealth quintile, education, and country dummies comprised the independent variables (appendix 2.7). Finally, we disaggregated these analyses by World Bank income group and WHO region (appendix 2.8).

All analyses were clustered at the primary sampling unit. We adjusted for stratification and applied sampling weights that accounted for unequal probability of selection, non-response, differences between the sample and the target population, missing survey weights, and missingness in covariates. Weights were constructed such that each country's weight corresponded to their 2015 population (appendix 2.6).

We did various sensitivity analyses to test the validity of our findings. We re-ran all analyses using equivalent weights whereby survey weights were rescaled such that each country contributed equally to the overall estimates. We also ran the hypertension analyses again but excluded data for India (the most populous contributor in the original analysis) to assess whether these data could skew our results. Finally, we applied three alternative sets of diagnostic testing criteria: the WHO PEN guidelines from the cardiovascular disease chapter for hypercholesterolaemia analyses (WHO, 2018a); the American College of Cardiology and American Heart Association guidelines (Arnett et al., 2019), which recommend diagnostic testing of all three risk factors for everyone aged 40-75 years (we excluded adults older than

75 years from this sensitivity analysis); and the WHO HEARTS guidelines (WHO, 2018a), which recommend hypertension testing for all adults, for the hypertension analyses. All analyses were done in STATA (version 17.0).

### Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

## 3.4. Results

Our final sample consisted of 57 surveys – 49 STEPS surveys and eight non-STEPS surveys – from 56 LMICs (Zanzibar had a survey separate from the rest of Tanzania because it has its own ministry of health and administers its health system largely independently of the rest of Tanzania). Maps of analysis samples (appendix 2.3) and survey characteristics by country (appendix 2.4) and individual countries' sampling methods are detailed in the appendix 2.5. Our sample comprised 994 185 individuals and was made up of three distinct cardiovascular risk factor samples (figure 3). The hypertension sample included 943 259 people in 55 LMICs, the diabetes sample included 225 707 people in 53 LMICs, and the hypercholesterolaemia sample included 250 573 people in 40 LMICs (table 8). The differences in sample size were due to varying availability of self-reported diagnostic testing information (appendix 2.4). Women accounted for 49·1% (95% CI 48·7–49·5) of the hypertension sample, 51·2% (50·7–51·6) of the diabetes sample, and 50·7% (50·3–51·1) of the hypercholesterolaemia sample (table 8).

19·1% (95% CI 18·5-19·8) of people in the hypertension sample, 23·8% (23·4–24·3) of people in the diabetes sample, and 27·4% (26·3–28·6) of people in the hypercholesterolaemia sample met the WHO PEN testing criteria (table 8). Self-reported testing was higher for hypertension (63·1% [62·4–63·8]) than for diabetes (28·6% [28·0–29·2]) or hypercholesterolaemia (29·8% [26·9–32·9]; table 8; appendix 2.9). Among the 170 810 people in the hypertension sample

*Table 8: Sociodemographic characteristics, by cardiovascular risk factor group*

| | Hypertension (n=943 259) | Diabetes (n=225 707) | Hypercholesterolemia (n =250 573) |
|---|---|---|---|
| Sex | | | |
| Female | 729,608 (49·1% [48.7-49.5]) | 132,743 (51·2% [50.7-51.6]) | 143,115 (50·7% [50.3-51.1]) |
| Male | 213,646 (50.9% [50.5-51.3]) | 92,960 (48.8% [48.4-49.3]) | 107,453 (49.3% [48.9-49.7]) |
| Age, years | | | |
| mean | 36·1 (12·8) | 38·8 (14·1) | 40·1 (14·9) |
| 18 - 39 y/o | 618,265 (62·8% [62.2-63.3]) | 106,599 (56·3% [55.7-56.9]) | 119,710 (52·6% [51.4-53.9]) |
| 40 - 64 y/o | 305,438 (34·3% [34.0-34.7]) | 106,871 (39·0% [38.4-39.6]) | 111,285 (40·8% [39.7-41.8]) |
| 65+ y/o | 19,556 (2·9% [2.6-3.2]) | 12,237 (4·7% [4.4-4.9]) | 19,578 (6·6% [6.0-7.2]) |
| Education | | | |
| Less than primary school | 242,841 (19·6% [19.2-20.0]) | 47,353 (19·5% [19.0-20.1]) | 35,517 (14·4% [13.2-15.6]) |
| Less than secondary school | 186,721 (24·7% [24.3-25.1]) | 74,971 (36·9% [36.1-37.6]) | 80,354 (35·8% [34.6-37.0]) |
| Secondary completed or higher | 510,522 (55·7% [55.1-56.3]) | 100,314 (43·6% [42.9-44.4]) | 132,609 (49·9% [48.5-51.3]) |
| Wealth quintile | | | |
| 1 (poorest) | 177,440 (17·9% [17.6-18.2]) | 36,680 (21·3% [20.6-22.1]) | 40,813 (18·4% [16.6-20.2]) |
| 2 | 176,015 (18·9% [18.6-19.2]) | 34,240 (21·3% [20.7-21.9]) | 42,010 (21·0% [20.0-22.0]) |
| 3 | 176,127 (19·8% [19.5-20.1]) | 32,955 (19·8% [19.2-20.5]) | 39,363 (20·1% [19.3-20.9]) |
| 4 | 175,374 (21·1% [20.7-21.4]) | 30,717 (19·0% [18.3-19.6]) | 37,328 (20·0% [19.1-20.8]) |
| 5 (richest) | 174,468 (22·3% [21.8-22.8]) | 28,790 (18·6% [17.5-19.7]) | 36,318 (20·6% [19.5-21.7]) |
| BMI | | | |
| mean | 23·4 (5·0) | 24·1 (5·3) | 24·4 (5·1) |
| < 18.5 (underweight) | 145,267 (13·5% [13.2-13.8]) | 17,583 (10·0% [9.6-10.4]) | 15,081 (8·7% [8.3-9.2]) |
| 18.5 to <25 (normal) | 528,744 56·2% [55.7-56.7]) | 106,161 (54·6% [54.1-55.1]) | 110,602 (51·7% [50.7-52.7]) |
| ≥ 25 (overweight or obese) | 269,248 30·3% [29.7-31.0]) | 101,963 (35·4% [34.9-35.9]) | 124,890 (39·6% [38.3-40.8]) |
| Cardiovascular risk factor prevalence* | | | |
| Hypertension | 182,984 21·6% (21.3-21.9) | 69,630 (26·6% [26.0-27.1]) | 77,402 (28·1% [27.4-28.9]) |
| Diabetes | 35,152 (5·3% (5.1-5.5) | 14,590 (7·0% [6.7-7.3]) | 12,289 (7·3% [6.7-8.1]) |
| Hypercholesterolemia | 11,277 (6·6% (6.2-7.1) | 11,172 (6·5% [6.2-6.9]) | 10,263 (7·0% [6.6-7.5]) |
| Testing recommended† | 170,810 (19·1% (18.5-19.8) | 76,225 (23·8% [23.4-24.3]) | 86,402 (27·4% [26.3-28.6]) |
| Tested for cardiovascular risk factor | 623,594 (63·1% (62.4-63.8) | 76,546 (28·6% [28.0-29.2]) | 96,016 (29·8% [26.9-32.9]) |

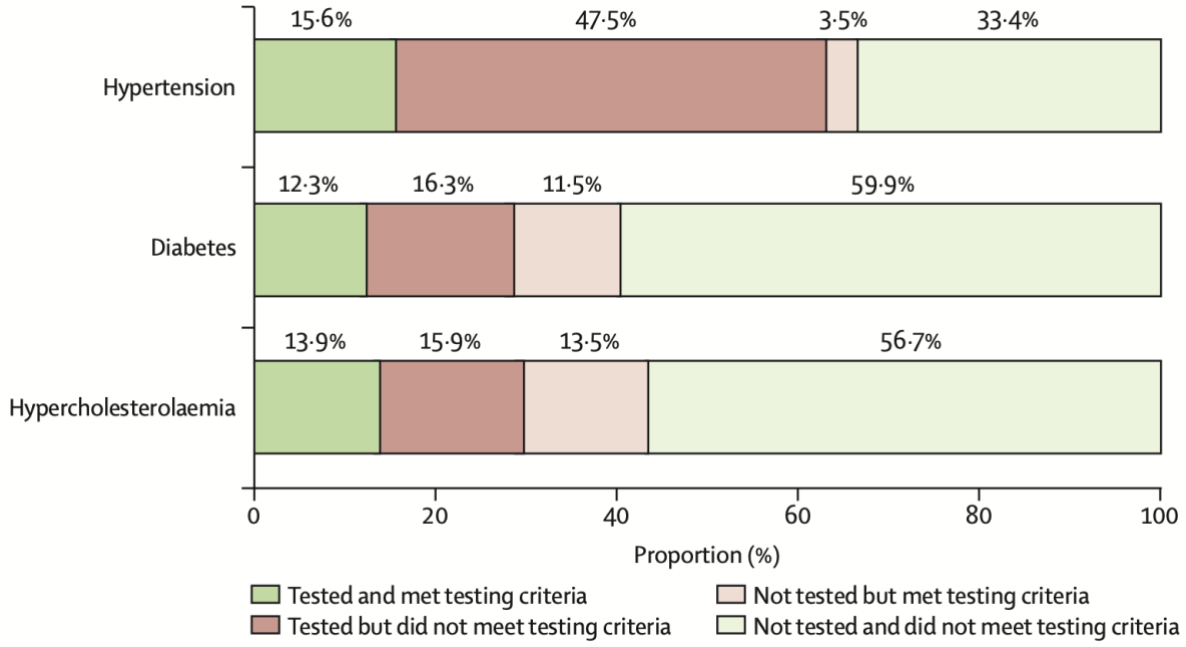
Notes: Data are n (% [95% CI]), where n is the number of observations, or mean (SD) for quantitative variables. The number of observations is unweighted, but for calculation of percentages, we accounted for sampling design, with survey weights rescaled such that countries' contributions corresponded to their population size.
*Defined according to results of the diagnostic testing that took place as part of the surveys.
†According to the WHO package of essential noncommunicable disease interventions for primary care guidelines, diagnostic testing was recommended for people with a BMI > 30 kg/$m^2$ and those older than 40 years with a BMI > 25 kg/$m^2$ (WHO, 2020).

who met the WHO PEN criteria for testing, 140 951 (78·6% [77·8–79·2]) had undergone diagnostic testing. The corresponding data for the other samples were 41 460 (44·9% [43·7–46·2]) of 77 044 in the diabetes sample and 48 571 (39·7% [37·1–42·4]) of 93 222 in the hypercholesterolaemia sample. Adherence to guidelines – ie the proportion of the samples

for which the fulfilment of testing criteria and testing status coincided (such that those in whom testing was not indicated did not receive a test and those in whom testing was indicated did receive on) – was recorded in 430 757 (49·0% [48·7–49·4]) of 943 259 people in the hypertension sample, 155 037 (72·2% [71·7–72·7]) of 225 707 people in the diabetes sample, and 158 477 (70.6% [69.7-71.4]) of 250573 people in the hypercholesterolaemia sample. In all three groups, most of the deviations for WHO PEN's testing recommendations were due to people being tested for a risk factor despite there being no indication for such testing (figure 3). Such testing was particularly pronounced in the hypertension group, in which 482 643 (47·5% [95% CI 47·1–47·9]) of 943 259 were tested for hypertension despite not meeting the WHO PEN criteria.

Diagnostic testing performance varied by individuals' sex, wealth quintile, and education level (figure 4). Women were 5 percentage points more likely than men to meet the WHO PEN criteria (ie, to have a BMI > 30 kg/m$^2$, or a BMI > 25 kg/m$^2$ while aged 40 years or older) in the hypertension sample, and approximately 10 percentage points more likely to do so in the diabetes and hypercholesterolaemia samples (appendix 2.9). Similarly, women were more

*Figure 3: Adherence to WHO PEN diagnostic testing recommendations, by cardiovascular risk factor*

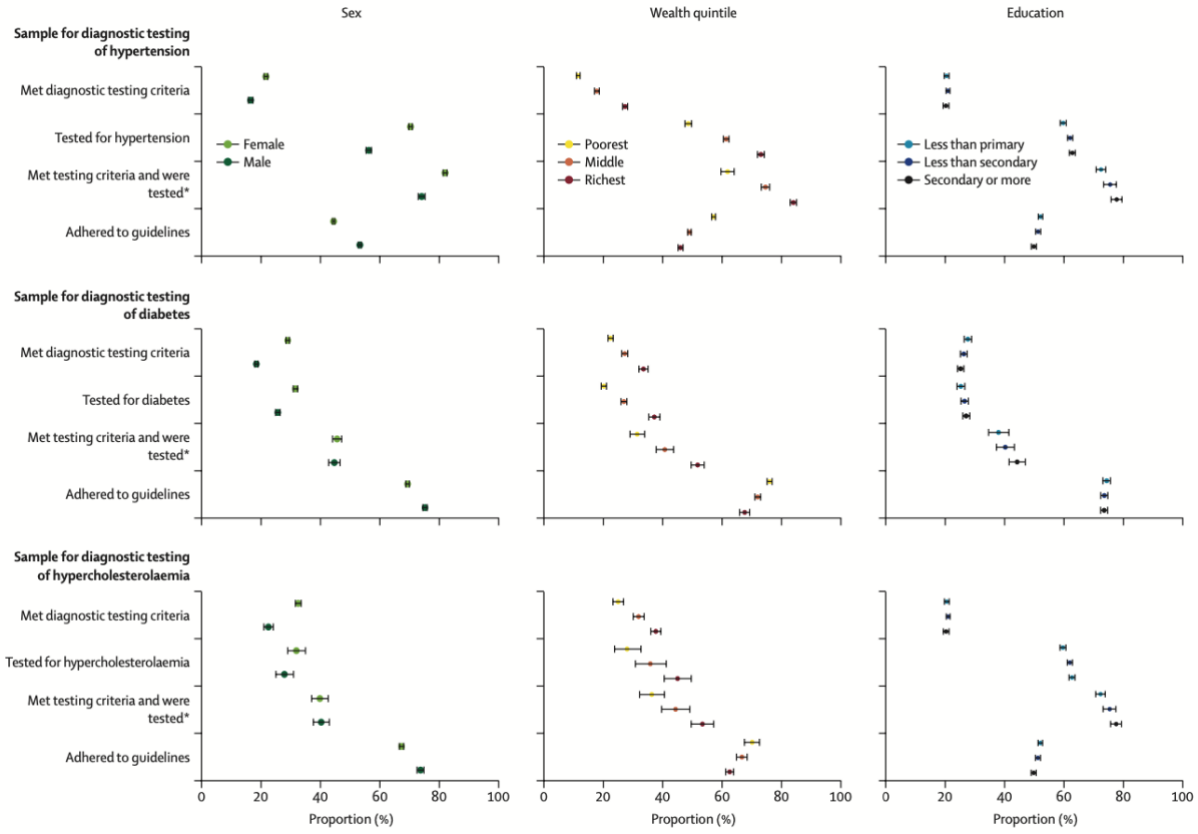



Note: According to the WHO PEN guidelines, diagnostic testing is recommended for all people with a BMI > 30 kg/m$^2$ and people older than 40 with a BMI > 25kg/m$^2$ (WHO, 2020).

likely than men to report having been tested for hypertension (p<0·0001) and diabetes (p<0·0001). Among people who met the WHO PEN criteria, the proportion who were tested was significantly higher among women than men in the hypertension sample (p<0·0001), but did not differ significantly between men and women for diagnostic testing of diabetes or hypercholesterolaemia (appendix 2.9). However, adherence to guideline criteria (ie being tested when testing was indicated, and not being tested when testing was not indicated) was significantly higher among men than among women for all three risk factors (figure 4). For hypertension testing, for example, recommendations were adhered to for 113 739 (53·3% [95% CI 52·8–53·8]) of 213 646 men and 317015 (44·4% [44·1–44·8]) of 729581 women. This discrepancy was due to a higher proportion of women than men being tested for hypertension despite such testing not being indicated by the WHO PEN criteria (appendix 2.10).

Among people who met the WHO PEN criteria, the proportion who had undergone testing

*Figure 4: Diagnostic testing performance by cardiovascular risk factor and sex, wealth, and education categories*




Note: Error bars represent 95% Cis. *The denominator for this population is the total number of people who met testing criteria.

was higher among individuals in the wealthiest quintile than those in the other quintiles (figure 3; appendix 2.9). Among people meeting the testing criteria for diabetes, 6 940 (51·8% [95% CI 49·6–54·0]) of 28 790 individuals in the richest quintile were tested compared with 4 850 (31·4% [29·0–33·9]) of 36 680 in the poorest quintile (appendix 2.9). Adherence to diagnostic testing criteria was lower in the richest than in the poorest quintiles for all three risk factors (figure 3; appendix 2.9). The worse adherence in the richest quintiles was driven by substantial proportions of people being tested for the risk factors despite not meeting the WHO PEN criteria (appendix 2.10). Individuals' education level was not strongly associated with their meeting of the WHO PEN criteria or their testing status (figure 3; appendix 2.9). Similarly, adherence to guidelines did not seem to differ relative to education for any of the risk factors (appendices 2.9 and 2.10).

Multivariable logistic regressions showed that women were more likely than men to get tested for all three risk factors, as were those with secondary education or more (compared with those with less than primary education) and those in higher wealth quintiles (compared with those in the lowest wealth quintile; table 9). The appendix presents the results disaggregated by World Bank Income Group and WHO World Region (appendix 2.11) as well as the results of all sensitivity analyses (appendices 2.12-2.17), which were largely similar to those of the main analyses.

## 3.5. Discussion

Our study has shown that, in LMICs, diagnostic testing performance for three major CVD risk factors is low and characterized by large deviations from testing recommendations. We further detected inequalities in access to diagnostic testing by sex and socioeconomic background. Overall, diagnostic testing for hypertension, diabetes, and hypercholesterolaemia — major risk factors for cardiovascular disease — followed socioeconomic gradients in LMICs. Among people who met the criteria in WHO PEN guidelines for diagnostic testing, more educated people and those in higher wealth quintiles were

*Table 9: Multivariable logistic regression analysis of associations between diagnostic testing and sociodemographic status among people who met the WHO PEN diagnostic testing criteria.*

| | Hypertension | | Diabetes | | Hypercholesterolemia | |
|---|---|---|---|---|---|---|
| | Odds Ratio (95% CI) | p value | Odds Ratio (95% CI) | p value | Odds Ratio (95% CI) | p value |
| **Sex** | | | | | | |
| Male | Ref | | Ref | | Ref | |
| Female | 1·86 (1·74-1·99) | <0·0001 | 1·21 (1·11-1·33) | <0·0001 | 1·25 (1·13-1·37) | <0·0001 |
| **Wealth quintile** | | | | | | |
| 1 (poorest) | Ref | | Ref | | Ref | |
| 2 | 1·23 (1·10-1·39) | 0·0004 | 1·12 (0·99-1·26) | 0·0683 | 1·11 (0·98-1·27) | 0·11 |
| 3 | 1·64 (1·47-1·83) | <0·0001 | 1·47 (1·27-1·71) | <0·0001 | 1·64 (1·44-1·87) | <0·0001 |
| 4 | 1·92 (1·71-2·16) | <0·0001 | 1·66 (1·40-1·97) | <0·0001 | 2·11 (1·83-2·44) | <0·0001 |
| 5 (richest) | 2·94 (2·59-3·34) | <0·0001 | 2·45 (2·13-2·82) | <0·0001 | 2·87 (2·48-3·33) | <0·0001 |
| **Education** | | | | | | |
| Less than primary | Ref | | Ref | | Ref | |
| Less than secondary | 1·15 (1·04-1·28) | 0·0047 | 1·03 (0·90-1·18) | 0·6902 | 0·91 (0·8-1·04) | 0·17 |
| Secondary or more | 1·39 (1·28-1·52) | <0·0001 | 1·33 (1·18-1·51) | <0·0001 | 1·36 (1·2-1·54) | <0·0001 |

Note: Models included a testing dummy as the dependent variable and sex, education category, wealth quintile, and country-fixed effects as independent variables. Survey weighting and clustering at the country level were accounted for.

significantly more likely to be tested for all three risk factors. Furthermore, people in higher wealth quintiles were more likely to be tested than those in the lowest wealth quintile even when they did not meet the diagnostic testing criteria, suggesting a suboptimal use of limited diagnostic resources.

On the one hand, a substantial number of individuals who met the WHO PEN criteria were not tested in each sample. On the other hand, we noted diagnostic testing in people who did not meet the WHO PEN criteria— particularly non-indicated testing for hypertension. Comparison of our findings for diagnostic testing coverage with those of previous studies (Geldsetzer et al., 2019; Manne-Goehler et al., 2019; Flood et al., 2021; Marcus et al., 2021) suggests that the wider adult population's access to risk factor testing for cardiovascular disease is lower than that of individuals with hypertension, diabetes, or hypercholesterolaemia. This could suggest that countries might be better at targeting diagnostic testing efforts to people at high risk of cardiovascular disease than our exercise of benchmarking testing performance against WHO PEN recommendations suggested.

Importantly, our study revealed major inequities in risk factor testing by sex, wealth, and education. Although significantly more women than men met the WHO PEN criteria, the proportions of women and men who had been tested (among all people who met the criteria) were equal in the diabetes and hypercholesterolaemia samples. Wealthier individuals were more likely to meet the WHO PEN criteria than those from lower wealth quintiles, and, among those who met the criteria, were more likely to be tested. Significant differences in diagnostic testing performance by educational attainment, however, were detected only via multivariable logistic regression that controlled for sex and wealth, in which individuals with higher educational attainment were more likely to be tested than those with lower educational attainment among people fulfilling diagnostic testing criteria. These differences in diagnostic testing access by sociodemographic characteristics were robust to various sensitivity analyses.

Our study had several limitations. First, we could not assess adherence to the third criterion of the WHO PEN guidelines — testing anyone with symptoms — because our sample did not include data for all possible symptoms at the time of respondents' diagnostic tests. Second, our measure of diagnostic testing (ie asking respondents if they had previously been tested) did not provide data for why and how often people had been tested. As a result, our weighted means might have been biased—eg we might have classed respondents who were tested because they had cardiovascular symptoms (despite not meeting the BMI or age criteria) as not meeting the WHO PEN criteria, leading to overestimation of poor adherence. Conversely, although pregnancy at the time of surveying was an exclusion criterion for this study, blood pressure might previously have been measured during pregnancy consultations, which could falsely inflate our estimated share of women fulfilling the WHO PEN criteria and being tested. Accordingly, our chosen measure of at least one previous diagnostic test cannot address the full range of testing rationales or assess testing quality but rather serves as an indicator of individuals having any testing access at all. Third, the WHO PEN diagnostic testing guidelines do not perfectly reflect the risk of developing cardiovascular disease (Teufel et al., 2021). BMI is an imperfect measure of obesity (Ashwell, Gunn and Gibson, 2012). Additionally, the guidelines we applied disregard other testing determinants, such as smoking or a history of premature cardiovascular disease, diabetes, or kidney disease in first-degree relatives, all of which are major causes of cardiovascular disease. Fourth, if countries followed national

diagnostic testing criteria different from those set out in the international WHO PEN guidelines, our assessments of guideline adherence could have been biased. Fifth, we did not assess access to diagnostic testing by rural versus urban location because many surveys did not record individuals' location category. Sixth, the 56 countries included in this study represent 39·9% of the world population and 47·6% of the population of LMICs, which means that our sample might not be representative of all LMICs. Seventh, even though the mean share of missing outcome data was low, a few countries had substantially higher proportions of individuals without diagnostic testing data (eg Iraq; appendix 2.4), which could have led to selection bias. Eighth, we used self-reported diagnostic testing data, which might be subject to recall or social desirability bias. Ninth, our logistic regression models were based on a timeless outcome variable and time-sensitive predictors, which could have led to people whose age and BMI were close to the guidelines' cutoffs only recently before the survey (rather than at the time of testing) fulfilling the WHO PEN criteria. Finally, we do not claim that the presented associations between sociodemographic characteristics and diagnostic testing performance are causal.

Notwithstanding these limitations, our analysis of testing of cardiovascular risk factors might be the best approximation of diagnostic testing access in LMICs to date. Health policy makers should aim to adhere to diagnostic testing guidelines and mitigate sociodemographic inequalities in testing access and uptake, given that diagnostic testing is the entry point to the care continuum for each of the presented risk factors and has major implications for downstream control of cardiovascular disease (Geldsetzer et al., 2019; Manne-Goehler et al., 2019; Flood et al., 2021; Marcus et al., 2021; NCD Risk Factor Collaboration, 2021).

# ESSAY THREE

## 4. The impact of patent expiry on statin consumption: A synthetic control analysis

Joint work with:

Sebastian Vollmer

### 4.1.Abstract

Healthcare systems optimize population health subject to a budget constraint. They choose which of the medically similar but differently priced drugs to grant market entry, prescribe and reimburse. Patent protection is a likely and important determinant in the optimization of drug consumption but remains underexplored methodologically and geographically.

We are the first to provide causal evidence and employ the synthetic control method. We use data on the consumption of two majorly consumed lipid-modifying agents, simvastatin and atorvastatin, in Germany, England and Sweden as an example.

We find a molecule's patent expiry increased its consumption but displaced therapeutically similar molecules. We argue that the main determinant of this substitution is the price competition by generic producers' market entry upon patent expiry which resulted in large cost savings. However, the substitution came at a cost of reduced healthcare quality due to a temporarily increased likelihood of side effects and worse compliance.

## 4.2.Introduction

High prices continue to hamper access to drugs important for prevention and treatment, in the Global South but also in high-income contexts (Cohen and Dupas, 2010; Herkert et al., 2019). For instance, the cost of a standard treatment for early-stage breast cancer amounts to the equivalent of 1.7 years of average annual wages in the USA, or 10 years in India or South Africa (WHO, 2018b). One important determinant of drug prices are patents. The patent protection of a drug grants its pharmaceutical producer the exclusive selling right, usually for twenty years (EU, 2023). Drugs' prices have been shown to be elevated during their patent protected period but then fall after their patent expires due to the subsequent entry of generic producers, introducing competition to the market (review by Vondeling et al., 2018).

The impact of patent expiry on drug consumption, however, remains less well understood.[14] Most studies have focused on the United States (Aitken, Berndt and Cutler, 2009; Berndt, Kyle and Ling, 2003; Duflos and Lichtenberg, 2012; Lakdawalla and Philipson; 2012) whose health system has been shown to be quite unique in comparison to the health systems of other high-income countries (Reibling, Ariaans and Wendt, 2019; Toth, 2020; Roser, 2020). They find the consumption of a drug to not increase after its patent expiry in the United States, as the overall decrease in a drug's price is accompanied by a decreased marketing effort for the drug by the producers of the original, branded drug (Aitken, Berndt and Cutler, 2009; Duflos and Lichtenberg, 2012; Lakdawalla and Philipson; 2012). While marketing expenditure has been shown to be a key determinant of drugs' consumption levels in the United States (eg Duflos and Lichtenberg, 2012; Lakdawalla and Philipson, 2012), this may not be true in other contexts like in European countries that are often characterized by a national or social health insurance system (Böhm et al., 2013). Studies that examine the impact of patent expiry on drug consumption outside of the United States find drug consumption to increase in England (Chapman, Fitzpatrick and Aladul, 2017) and a region in Italy (Fiorentini, Bruni and Mammi, 2022) or to remain unchanged in Sweden and Japan (Imai, Fushimi and Sundell, 2018).

[14] Instead of considering the impact on the overall molecule consumption, many studies have rather focused on changes in market shares, examining patent expiries' impact on the relative consumption of brand and generic products (Bae, 1997; Boersma et al. 2005; Castanheira, Ornaghi and Siotis, 2019; Clarke and Fitzgerald, 2010; Fischer and Stargardt, 2016; Grabowski, Long and Mortimer, 2014; Selvaraj, Farooqui and Mehta, 2019).

Prior research on the impact of patent expiry on drug consumption not only stems from a small number of countries but also relies on descriptive or regression-based methodologies. While Chapman, Fitzpatrick and Aladul (2017) and Fiorentini, Bruni and Mammi (2022) employ an interrupted time series (ITS) design, all other studies use descriptive, regression or time-series analyses only (Aitken, Berndt and Cutler, 2009; Aitken et al., 2013; Duflos and Lichtenberg, 2012; Lakdawalla and Philipson; 2012). The quality of the studies additionally varies in the data frequency and time periods (see appendix 3.1).

In this paper, we determine the impact of patent expiry on the consumption of two major lipid-modifying agents, simvastatin and atorvastatin, using monthly data from three European countries: Germany, England and Sweden.[15] Lipid-modifying agents are drugs prescribed to lower the cholesterol levels in the blood and thereby prevent the development of cardiovascular disease (NHS, 2023a), responsible for a third of all global deaths in 2019 (IHME, 2023).

We employ a novel, quasi-experimental strategy called the synthetic control method (SCM), as introduced by Abadie and Gardeazabal (2003) and summarized by Abadie (2021). We construct synthetic controls of simvastatin and atorvastatin by using a combination of other drugs acting on the cardiovascular system, which tell us what simvastatin and atorvastatin's consumption would have looked like had their patents not expired. In contrast to traditional regression or time-series analyses, it is particularly well-suited for examining the impact of large-scale, one-off interventions without requiring large samples or many observations of the intervention of interest (Abadie, 2021). While this relatively recent causal inference method has already been applied to many academic fields, including health economics (eg Bauhoff, 2014; Kreif et al., 2016; Mitze et al., 2020), to the best of our knowledge it has never been used to assess the impact of patent expiry on drug prices, consumption or health outcomes. Additionally, we replicate the results obtained from employing SCM with an interrupted time

[15] We additionally provide descriptive evidence from a further 13 countries using yearly data.

series analysis, an econometrically speaking weaker method[16] employed by other studies already (Chapman, Fitzpatrick and Aladul, 2017; Fiorentini, Bruni and Mammi, 2022).

We find simvastatin and atorvastatin consumption to increase in response to their patent expirations in Germany, England and Sweden. For instance, atorvastatin's consumption increased by 53% within a year after its patent expiry in May 2012 in England, by 75% in Sweden and by 788% in Germany. Simultaneously, we observe a negative impact on the consumption of other lipid-modifying agents due to displacement. We argue that the price reduction in response to patent expiry is the main driver of these results. This high sensitivity to price changes implies that these health systems have a high elasticity of substitution. On the upside, this results in substantial cost savings. Switching from atorvastatin to generic simvastatin, for example, resulted in €220 million savings for Germany in only nine months in 2003, while consistent substitution could have saved an additional €107 million (Klose and Schwabe, 2004b) and could have saved £2 billion between 2001 and 2006 in the UK (Moon and Bogle, 2006).

On the downside, however, the predominant prescription of simvastatin between 2003 and 2012 in Germany came at a health cost. When simvastatin lost its patent protection and became available generically in May 2003, it pushed out atorvastatin and became the predominantly prescribed statin until atorvastatin's patent expired in May 2012. While we lack monthly data dating back to before 2003 for England and Sweden, we nevertheless observe a predominant use of simvastatin prior to 2012 in both countries as well. The predominant consumption of simvastatin in that time period first meant an increased risk of developing rhabdomyolysis, a severe form of muscle pain that is the clinically most relevant side effect of lipid-modifying agents (Law and Rudnicka, 2006) as the likelihood of suffering from rhabdomyolysis is higher for simvastatin than atorvastatin (Egan and Colman, 2011). Second, adherence to the lipid-modifying agent regiment is more difficult with simvastatin than atorvastatin as simvastatin's therapeutic effect is sensitive to being taken punctually at night (Schachter, 2005). Lastly, it has been shown that switching patients away from an atorvastatin

[16] Cook and Campbell (1979) explain the shortcomings of an ITS design in cases of interventions having a gradual and / or delayed impact.

regiment increases non-adherence and discontinuation (Pettersson et al., 2012; Stargardt, 2010). In conclusion, all three health systems seem to have prioritized cost saving over healthcare quality between 2003 and 2012.

This paper is the first to employ the synthetic control method to determine the impact of patent expiry on drug consumption in Germany, England and Sweden using monthly data. We find that the patent expiry of individual lipid-modifying agents significantly increases their consumption by substituting away from more expensive molecules. The health systems' price sensitivity results in large cost savings but also an increased rhabdoymyolysis prevalence between 2003 and 2012.

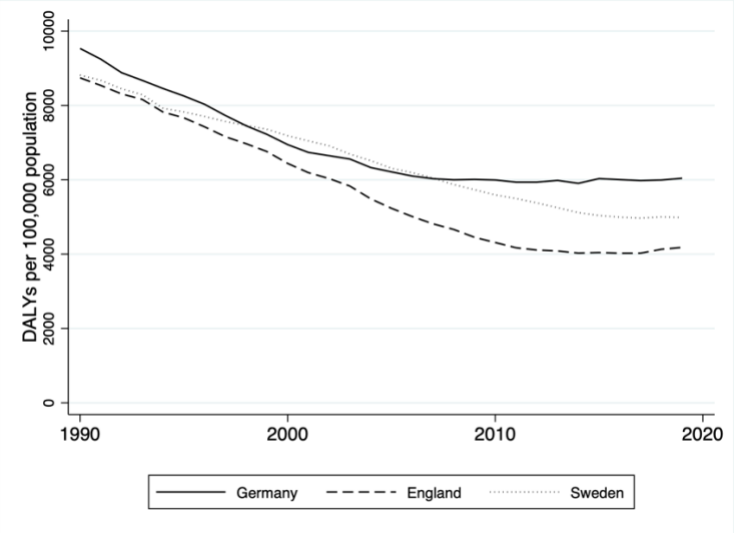
This paper proceeds as follows. Section 4.3 provides an overview of cardiovascular disease and the role of lipid-modifying agents in its prevention and describes the institutions and policies of the German, English and Swedish healthcare systems regulating pharmaceutical drug consumption. Section 4.4 proceeds to describe the data and our estimation framework. We discuss our results and robustness checks in section 4.5 before concluding in section 4.6.

## 4.3. Background

### Cardiovascular disease and its prevention with lipid-modifying agents

Cardiovascular disease (CVD) is a term used for any blood vessel or heart diseases or events such as strokes or heart attacks (NHS, 2023a) and is responsible for a third of all global deaths in 2019 (IHME, 2023). The rising disease burden of cardiovascular disease over recent years was driven by an increasing CVD burden in low- and middle-income countries (IHME, 2023; Peiris et al., 2021). In Germany, England and Sweden, on the other hand, the disease burden of cardiovascular disease has been decreasing (Figure 5). Figure 5 displays the burden of disease in the units of disability adjusted life years (DALYs), which corresponds to the number of years of healthy life lost, ie the number of life years lost due to premature mortality plus the number of years lived with a disease or disability. In the case of cardiovascular disease in Germany, England and Sweden, the decreasing disease burden is driven by a decrease in the

*Figure 5: The burden of cardiovascular disease in Germany, England and Sweden*




Source: IHME (2023)

number of premature deaths due to CVD, while the years lived with CVD have remained relatively stable since 1990 (see appendix 3.2).

One major risk factor of cardiovascular disease are high cholesterol levels in the blood over longer time periods, as they cause a plaque build-up in the arteries which can eventually lead to CVD events (CDC, 2023). Cholesterol can be lowered with life-style changes such as "a heart-healthy diet, regular exercise habits, avoidance of tobacco products, and maintenance of a healthy weight" (Stone et al., 2013) as well as taking medications. Among the lipid-modifying agents prescribed to lower cholesterol, a group of molecules known as statins make up the majority of lipid-modifying agent consumption (Klose and Schwabe, 2020). Overall, statins are considered safe as side effects occur very rarely.[17] The most clinically relevant side effect of statins is muscle pain, known as myopathy (Mach et al., 2020). Its most severe form, rhabdomyolysis, is a rapid muscle breakdown that can lead to renal failure and death and occurs approximately 1-3 times per 100,00 patient-years (Law and Rudnicka, 2006). The likelihood of muscle pain side effects generally rises with a statin's dosage (SEARCH Collaborative Group, 2010).

[17] 10mg of simvastatin, a statin molecule, was even made available without prescription over the counter in the United Kingdom in 2004 (Simvastatin over the counter, 2005).

Statins have been available since the 1990s. Lovastatin was the first molecule that entered the German market in 1989 while simvastatin followed a year later in 1990 (Klose and Schwabe, 2004b). A total of eight statins were developed and sold, seven of which are still available today.[18] In 1994, the first longitudinal study that proved statins' effectiveness in CVD prevention was published (Scandinavian Simvastatin Survival Study Group, 1994). In Europe, the patent of the first statin, simvastatin, expired in May 2003.[19] The second major statin molecule, atorvastatin, lost its European patent protection in May 2012 (Chapman, Fitzpatrick and Aladul, 2017).

Lipid-modifying agent consumption has dramatically increased since 2000 due a large increase in statin consumption (own data presented below; Blais et al., 2021). In fact, statins are one of the most used groups of drugs worldwide with approximately 173 million people consuming statins in 2018 (Blais et al., 2021). The two statin molecules simvastatin and atorvastatin make up approximately 90% of statin consumption.[20] They both frequently feature among the most prescribed drugs globally (Haqqi, 2023)[21] and atorvastatin is estimated to be the drug with the highest lifetime sales ever ($150 billion by Q3 of 2017; Brumley, 2017).

Simvastatin and atorvastatin are therapeutically reasonably good substitutes for each other which is also indicated by most medical guidelines or reimbursement policies recommending not a specific statin molecule but simply the group of statins (see next subsection). Even though both simvastatin and atorvastatin should be taken once daily, differences in ease of use, potency and side effects remain. Simvastatin should be consistently taken at night as its elimination half-life amounts to only two hours (Schachter, 2005) and the body's own cholesterol synthesis is most active at night (Jones and Schoeller, 1990). Atorvastatin's half-life

[18] The pharmaceutical company Bayer withdrew its statin cerivastatin globally on 8 August 2001 due to a significantly higher likelihood of the side effect rhabdomyolysis occurring, with 31 lethal occurrences (Marwick, 2003).
[19] Interestingly, in the United States, simvastatin's patent only expired in June 2006 (Aitken, Berndt and Cutler, 2009).
[20] Example figures based on author's own calculations: Simvastatin and atorvastatin were jointly responsible for 94.6%, 89.4% and 90.8% of all lipid-modifying agents consumption in England (2019), Germany and Sweden (both 2020), respectively.
[21] In 2023, simvastatin ranked twelfth with 36.6 million prescriptions and atorvastatin ranked first with 114.5 million prescriptions (Haqqi, 2023).

of 14 hours means it can be taken at any time of the day (Schachter, 2005). Their potency also differs as atorvastatin is able to effect approximately the same percentage reduction in cholesterol with only half the number of milligram as simvastatin (Weng et al., 2010). As the occurrence of rhabdoymyolysis is linked to the dosage (explained above), this means an occurrence of this side effect is more likely with simvastatin than atorvastatin (Egan and Colman, 2011). For this reason, guidelines actively recommend against the use of 80mg simvastatin (US Food and Drug Administration, 2011; NICE, 2014).

The efficacy of both simvastatin and atorvastatin depends on patients' compliance to the prescribed drug regiment which has been shown to deteriorate when switching the prescribed lipid-modifying agent. Stargardt (2010) found that when patients on an atorvastatin regiment were switched to simvastatin in response to a policy change in reimbursement rules in 2005 in Germany, they were more likely to not adhere to or discontinue their simvastatin treatment. Pettersson et al. (2012) similarly found higher non-adherence and discontinuations among patients who were switched away from an atorvastatin or rosuvastatin treatment in Sweden, also in response to a reimbursement policy change.

## The German healthcare system

There are two types of health insurance funds in Germany – public and private – with 90% of the population being insured with a public health insurance fund (GKV-Spitzenverband, 2023a). Since we will analyze claims data from public health insurance funds in Germany later, we proceed by focusing on public health insurance funds. Since the German Social Code mandates that healthcare should correspond to current, best medical standards (Book V, §2, section 1, sentence 3), public health insurance funds reimburse all prescription drugs. Patients only have to pay a contribution of 10% on prescription drugs ("co-payment") with the additional constraint that the co-payment is always at least 5€ and at most 10€.[22]

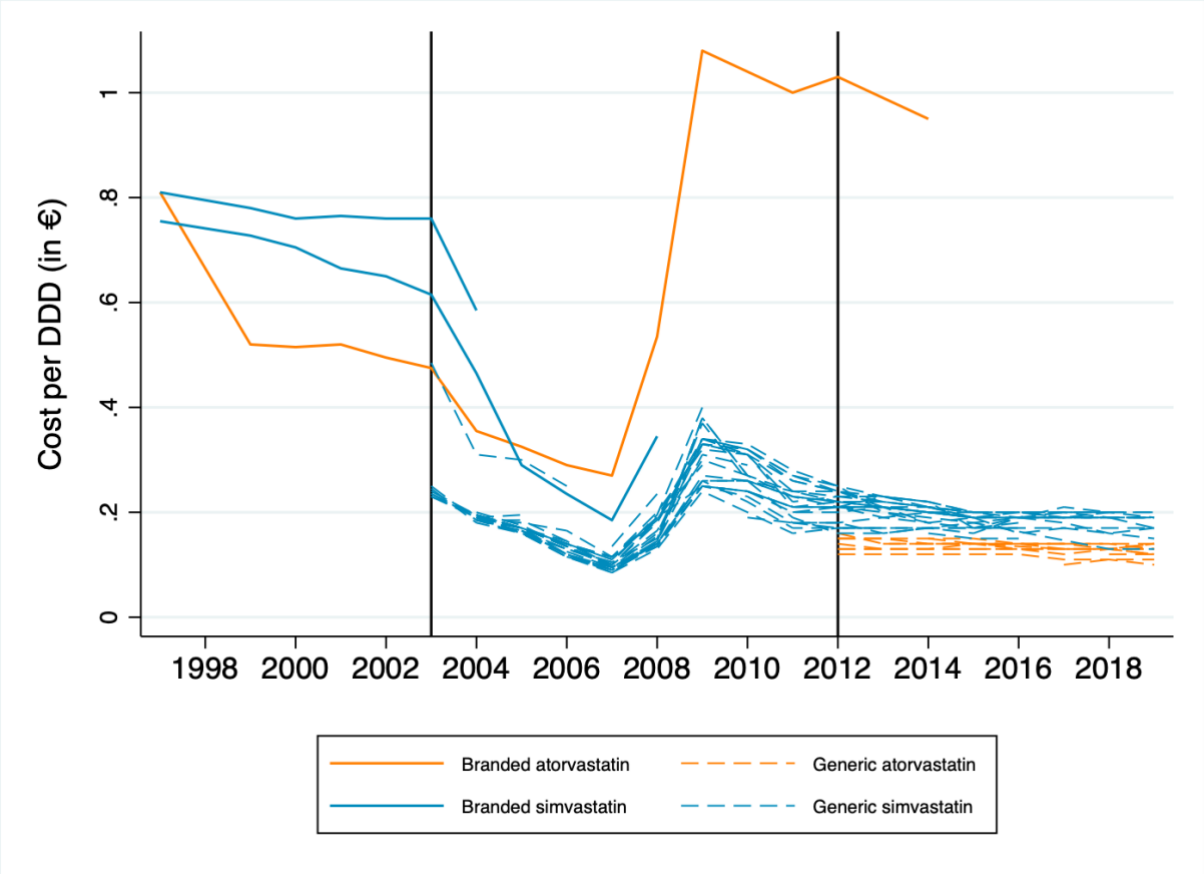
[22] There are some exceptions: Under-18-year-olds are exempt from co-payments. There is a co-payment ceiling so that patients do not have to pay more than 2% of their gross income. Patients with a "severe" chronic disease do not have to pay more than 1% of their gross income. When a drug price falls below 30% of the reference price, copayments are waived (Bundesgesundheitsministerium, 2023b).

Germany allows producers to freely set prices for brand and generic drugs, which can be changed twice per month (IFA, 2023). Prior to 2005, public health insurance funds reimbursed the statin prices as set by the pharmaceutical producers. From January 2005 onwards, however, statins were included in a reference price policy for therapeutically similar molecules. Germany had introduced a reference price system in 1989 already which regulated the maximum reimbursement amount by all public health insurance funds – the reference price. If a molecule's selling price exceeded this reference price, patients would have to pay the difference out of pocket. Alternatively, patients could discuss alternative regiments that involved molecules with prices below the reference price with their GP. From 1 January 2005 onwards, however, reference prices could now be introduced for groups of molecules and lipid-modifying agents were among the first molecules to be grouped together due to their therapeutic and chemical similarity and assigned a single reference price.[23] Reference prices are set and regularly updated by the National Association of Statutory Health Insurance Funds ("Spitzenverband Bund der Krankenkassen") according to the rules set out in the German Social Code (Book V, § 35 sections 1, 3 and 5). They are required to announce changes to their reference prices on their website (GKV-Spitzenverband, 2023b). We explain the process of setting reference prices and display those for statins in the appendix 3.3. Since 2006, it is additionally possible for health insurance funds to negotiate discount prices with pharmaceutical companies which means that the reference prices should be thought of as a price ceiling.

Figure 6 displays the average price of simvastatin and atorvastatin per DDD from each pharmaceutical producer available since 1997. Data are taken from the yearly pharmaceutical prescriptions report (Klose and Schwabe, 1999, 2001-2003, 2004a, 2004b, 2006, 2007, 2008a, 2008b, 2010-2020). Prior to simvastatin's patent expiry in May 2003, it shows that the price of atorvastatin (0.50€/DDD in 2002) was smaller than that of simvastatin (0.65€-0.76€/DDD in 2002). Generic simvastatin, in turn was much cheaper than branded atorvastatin or simvastatin and eventually drove branded simvastatin completely out of the market. A similar

[23] The grouping together of molecules in reference price groups is done by the Federal Joint Committee ("Gemeinsamer Bundesausschuss") which is a joint organization of all doctors, dentists, hospitals and health insurances in Germany. The grouping together of molecules due to chemical and therapeutic similarity, as is the case for lipid-modifying agents, is referred to as grouping type 2.

*Figure 6: Prices of branded and generic simvastatin and atorvastatin in Germany*




Note: Vertical lines correspond to patent expiry of simvastatin (2003) and atorvastatin (2012). There were4 two branded simvastatin products available prior to simvastatin's patent expiry already. For comparability across time, we converted early prices from Deutsche Mark to Euro and adjusted data pre-2009 such that all data points apply the current DDD definitions. Source: Klose and Schwabe, 1999; 2001; 2002; 2003; 2004a; 2004b; 2006; 2007; 2008a; 2008b; 2009; 2010; 2011; 2012; 2013; 2014; 2015; 2016; 2017; 2018; 2019; 2020.

pattern can be observed upon atorvastatin's patent expiry in May 2012, with generic atorvastatin driving out branded atorvastatin. It can also be seen that generic atorvastatin had a lower per-DDD price than generic simvastatin. We additionally plot individual package prices relative to the reference price in select years in appendix 3.3 (appendix figure A3.16).[24] These price dynamics are in line with what the literature has found, namely that prices typically decrease after a molecule's patent expires due to the price competition by generic producers entering the market (Vondeling et al., 2018).

[24] While simvastatin's price was on average smaller than the reference price in all years, this was true for atorvastatin only after 2012. Beforehand, atorvastatin was, on average, more expensive than the reference price.

Another mechanism to contain medical spending has been in effect since May 2006 according to which health insurance funds and doctors are required by law to negotiate yearly caps on drug expenditures.[25] They declare preferred substances within high-volume drug classes and set targets on how much of the expenditure within a drug group has to come from the preferred substance (buzer.de, 2023). Technically, these trickle down to the General Practitioner (GP) level as quotas which mandate how much of a drug class's expenditure share – like that of lipid-modifying agents – should come from the preferred substance(s). However, these rules are rarely enforced (McGuire and Bauhoff, 2011) which can also be seen in our data. Lipid-modifying agents were grouped together and simvastatin determined the preferred substance in 2007, pravastatin was added as a second preferred substance in 2013, and atorvastatin became a third preferred substance only in 2022 (Kassenärztliche Bundesvereinigung, 2007; 2011; 2021). We present the target expenditure shares in appendix 3.4.

## The English healthcare system

The National Health Service (NHS) provides free healthcare to all residents of England and, in fact, the remainder of the United Kingdom (UK) as well and is financed through taxes. GP and hospital visits are provided free (except for dental or optical care) of charge but patients face co-payments for prescription drugs (Office for Health Improvement and Disparities, 2023). Currently, patients must pay £9.65 per prescription drug except for contraceptives or drugs received as a hospital inpatient (NHS, 2023b).[26]

Prices for branded drugs are regulated through the Pharmaceutical Price Regulation Scheme (PPRS) and for generic drugs through the Drug Tariff (Houses of Parliament, 2010). The PPRS is a non-binding agreement between the UK's health ministry and the pharmaceutical industry

[25] §84 of the German Social Code Book V requires the National Association of Statutory Health Insurance Funds ("GKV-Spitzenverband") and the National Association of Statutory Health Insurance Physicians ("Kassenärztliche Bundesvereinigung", KBV) to set these targets for a new year by 30 September of the previous year. The targets are then broken down for each regional Association of Statutory Health Insurance Physicians.

[26] Many individuals can get exemptions from these co-payment charges, though. For example: individuals who are aged over 60 years or under 16 years, individuals who are aged between 16 and 18 and are in full-time education, women who are pregnant or have had a child in the previous 12 months, individuals with certain disabilities, individuals on income support or on jobseeker's allowance (NHS, 2023b).

which is renegotiated every five years (McGuire and Bauhoff, 2011). It caps by how much NHS spending on branded drugs grows over the next five-year cycle as well as how much profit pharmaceutical companies are allowed to make from their sales to the NHS. As these latter profit caps are at the company and not pharmaceutical product level, companies are free to set their prices such that the profitability of their individual products may vary provided overall profits made by NHS sales are kept below the cap. Any excess profits are to be paid back to the health ministry which rarely occurs in practice as excess profits are offset through price cuts (Houses of Parliament, 2010). The PPRS used to also have the power to mandate price cuts but with its last negotiation rounds in 2018 switched to a payment scheme where pharmaceutical producers pay a percentage of sales back to the health ministry (Department of Health & Social Care, 2023).

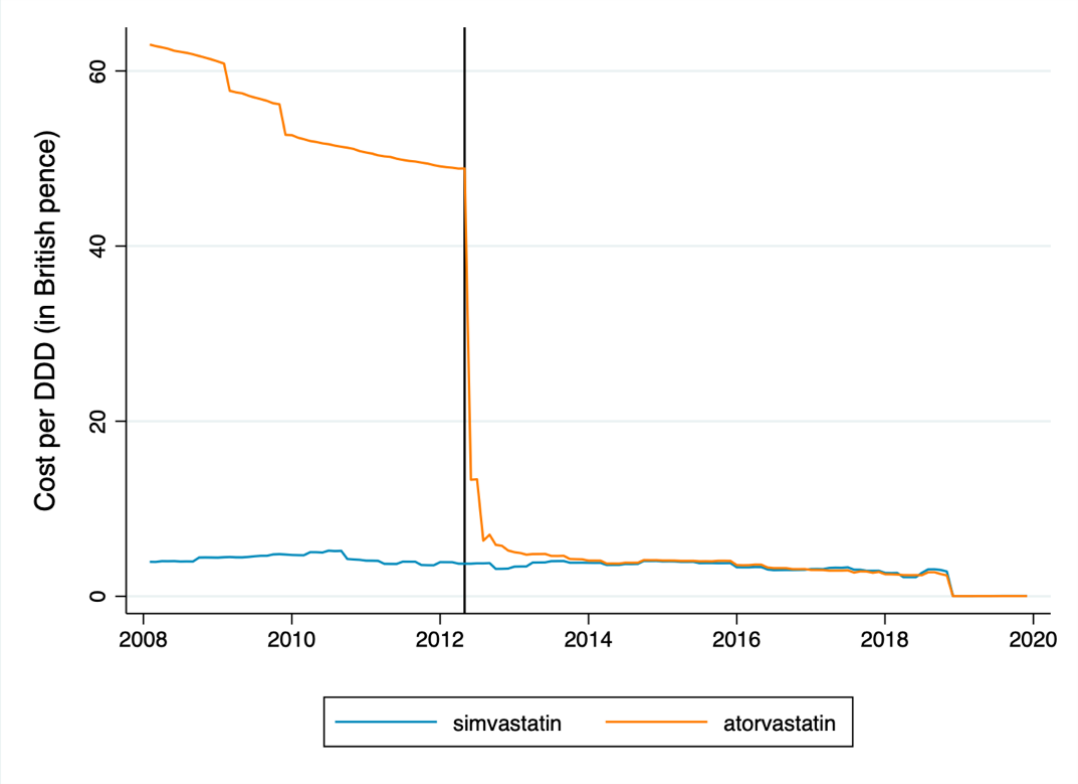
The Drug Tariff, on the other hand, is published each month by the health ministry and determines how much pharmacies are paid for each generic dispensed for a prescription (National Audit Office, 2018). These reimbursement rates are based on "volume-weighted average ex-factory prices across available generics in the UK" (McGuire and Bauhoff, 2011) and not on pharmacies actual operating costs. This resulted in substantial price decreases for generics in the UK (Office of Fair Trading, 2007).

However, the mechanisms to incentivize switching to cheaper molecules within drug classes or to cheaper generics are weak in the UK (McGuire and Bauhoff, 2011). While there are prescription targets arranged between Primary Care Organizations and GP practices, they are not made with individual GPs (McGuire and Bauhoff, 2011). GPs are encouraged but not mandated to prescribe generics when available and research has found GPs to be unaware of prices and price changes of drugs in the UK (Office for Fair Trading, 2007). Additionally, there are no rules for pharmacies to dispense generics when GPs prescribe a branded drug (McGuire and Bauhoff, 2011).

All lipid-modifying agents are prescription drugs and reimbursable in the UK. McGuire and Bauhoff (2011) find that simvastatin's patent expiry in May 2003 led to significant substitution away from branded to generic simvastatin but very little substitution away from atorvastatin. Chapman, Fitzpatrick and Aladul (2017) similarly found price to be a major determinant of

lipid-modifying agent consumption in England. Lastly, this is also reflected in our consumption data displayed later in our results section as well as in Figure 7 which displays the per DDD cost that simvastatin and atorvastatin were responsible for in the English health budget. It shows that the per DDD atorvastatin cost dropped substantially in response to its patent expiry in May 2012.

*Figure 7: Per DDD cost of simvastatin and atorvastatin in England*




Note: Vertical line corresponds to atorvastatin's patent expiry. Data source and cleaning identical to that of molecule consumption as outlined in appendix 3.7.

## The Swedish healthcare system

All residents of Sweden are automatically covered by its national health service which is provided in a decentralized manner by counties and municipalities and financed through taxes (TLV, 2017). Patients face co-payment fees for in- or outpatient visits and drug treatments which are capped. Within a 12-month period, patients must pay the full amount of any prescribed drugs up to SEK 1 300 (approx. €110), after which partial subsidies apply until the patient has paid SEK 2 600 (approx. €220) with drugs being fully subsidized thereafter (TLV, 2023).

Pharmaceutical producers are free to set their own prices. If products are subject to prescriptions and should be reimbursed by the healthcare system, however, pharmaceutical producers must apply to the Dental and Pharmaceutical Benefits Agency (Swedish acronym: TLV). In their application, producers must state their product's price and enclose health economic documentation (TLV, 2023). The TLV determines whether the product will be reimbursed ('included in the Pharmaceutical Benefits Scheme') based on the three principles of human value (ie no discrimination when making reimbursement decisions), need and solidarity (ie those in greatest need of medical treatment are prioritized when making reimbursement decisions) and cost-effectiveness (Ministry of Health and Social Affairs, 1982). The TLV and this reimbursement system were established in 2002 and all drugs previously reimbursed were to be reviewed according to these principles until 2010 (Pettersson et al., 2012).

As lipid-modifying agents are also subject to prescriptions in Sweden, they were part of these policy changes. Prior to the review of lipid-modifying agents in 2009, the Swedish health service was reimbursing all lipid-modifying agents. From June 2009 onwards, however, reimbursement rules changed depending on the prices for which individual lipid-modifying agents were available (TLV, 2009): Generic simvastatin (and pravastatin) continued to be fully reimbursed, whereas the reimbursement of atorvastatin (and rosuvastatin) was now restricted. 10mg atorvastatin was completely excluded from reimbursement but higher dosages were reimbursable if patients had tried but not met their target blood lipid levels with higher dosages of generic simvastatin, first (Pettersson et al., 2012).[27] However, pharmaceutical producers of atorvastatin or other drugs under the restricted reimbursement rule were allowed to apply for reimbursement status and would be granted a reimbursement status if considered cost-effective in comparison to generic simvastatin (TLV, 2009).

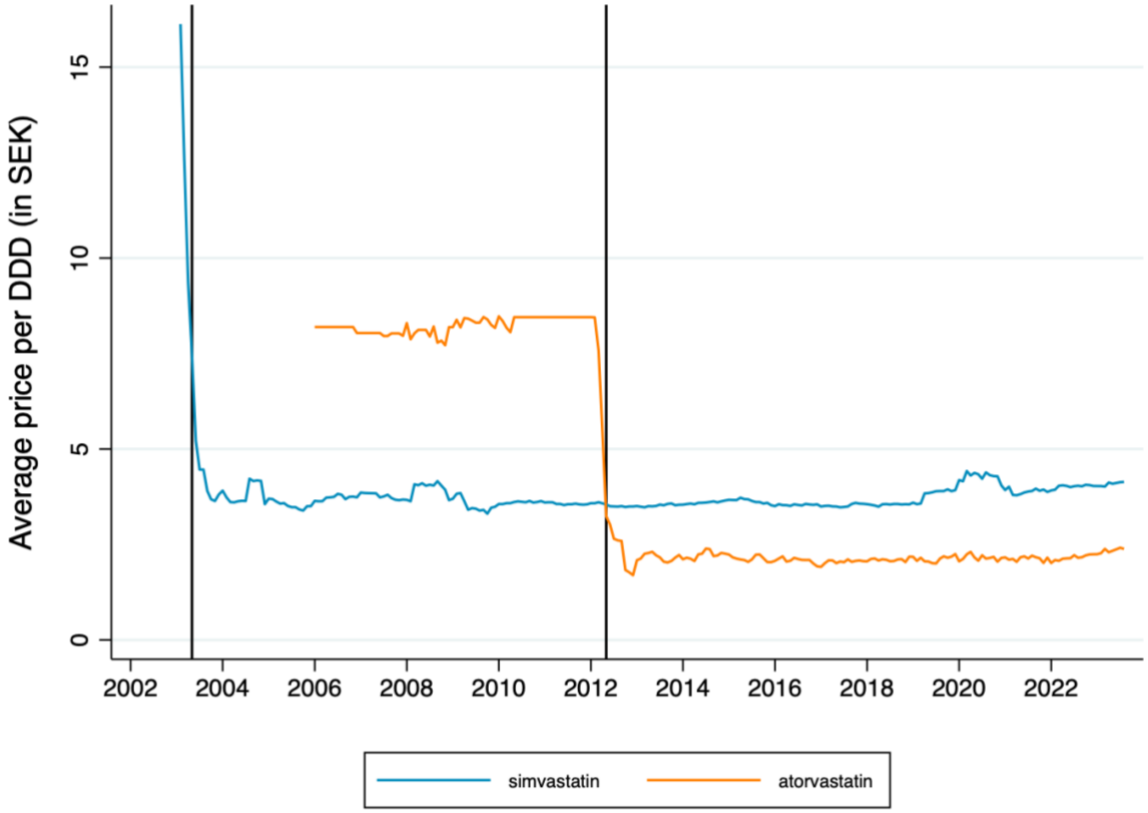
Another mechanism indirectly influencing prices of lipid-modifying agents in Sweden is a monthly auction-based system at the pharmacy level. Since its introduction in 2002, the TLV

[27] Similarly, 5mg rosuvastatin was also excluded from reimbursement, as were fluvastatin, pravastatin, cholestyramine and branded simvastatin (Pettersson et al., 2012).

decides which product in each molecule's package-size group should be dispensed depending on which package has the lowest retail price (TLV, 2017). Pharmacies are obliged to sell this cheapest package of the drug prescribed by the GP (unless the GP objects and provides a reason) (TLV, 2023).

In Figure 8, we display the average price per DDD for simvastatin and atorvastatin. It can be seen clearly that the dates of their patent expiries occur simultaneously with large price drops. Additionally, atorvastatin's average per DDD price falls below that of simvastatin in response to atorvastatin's patent expiry in May 2012. These results are similar to the price drops observed in Germany and England in response to patent expiry and are in line with what the literature has found (Vondeling et al., 2018).

*Figure 8: Prices of simvastatin and atorvastatin in Sweden*




Note: Data sources and cleaning procedures are outlined in appendix 3.5. Vertical lines correspond to simvastatin (May 2003) and atorvastatin's (May 2012) patent expiry dates.

## 4.4. Methods

### Data

We utilize monthly data on all lipid-modifying agent molecules from Germany, England and Sweden for our analyses but also present descriptive results from 13 additional countries using yearly data. We provide an overview of the data sources and the available time frames in appendix 3.6. We conduct all data cleaning, harmonization and analysis in the statistics software package STATA 17.0

Data for Germany were generated using the database of the German Institute for Drug Use Evaluation (DAPI) containing anonymous dispensing data of community pharmacies claimed to the public health insurance funds. The English and Swedish data were publicly available through the website of the National Health Service Business Services Authority (NHSBSA, 2021) and the National Board of Health and Welfare, a government agency of the Swedish Ministry of Health (Socialstyrelsen, 2023), respectively.

We harmonize our data using the World Health Organization's Collaborating Centre for Drug Statistics (WHOCC, 2018) methodology: The WHOCC classifies each medical molecule in a five-tiered Anatomical Therapeutical Chemical (ATC) system where lipid-lowering drugs can be found in group C (Cardiovascular system) subgroup 10 (Lipid-modifying agents). For example, simvastatin is referred to with the ATC code C10AA01. We assign each molecule in our data the corresponding ATC code and later use the ATC level 2 subgroups, C01, C02, etc., to determine our control groups for the synthetic control method.

Where possible, we additionally standardize the dosages of each molecule using the WHOCC's Defined Daily Dosage (DDD) definition which approximately represents a molecule's average daily dosage in adults (WHOCC, 2018).[28] The DDD of simvastatin, for instance, is 30 mg. As the definitions of DDD changed for some of our lipid-lowering agents' molecules in 2009, we use the current DDD definitions used from 2009 onwards and convert DDD pre-2009 to the current

[28] The definition of DDD was changed for some molecules in our sample in 2009. We use the current DDD definitions used from 2009 onwards and convert DDD pre-2009 to the current definitions where necessary.

definitions where necessary. Data for Germany are calculated by aggregating the number of milligrams over all dosages and package sizes for each molecule and month and then converting the number of milligrams to DDD for each molecule-month combination. We follow a similar process for England which we explain in appendix 3.7. The Swedish data unfortunately do not provide the granularity required for a conversion to DDD units. It only reports the number of dispensings from which we could not deduce the dosages or package sizes. We argue this is an imperfect but comparable measure as we show that using the number of prescriptions instead of DDD per month in the case of England yields similar results. We provide a visual comparison of prescription and DDD data and run our analyses using English prescription data as a robustness check in the appendix 3.14.

Lastly, we divide all DDD and dispensings per month by 1,000 population. For Germany, we use the number of patients insured with public health insurance funds (Bundesgesundheitsministerium, 2023a) which are available monthly from January 2003 onwards. For the years 2000-2002, we use the yearly numbers for all 12 months per year as the number of people insured with public health insurance funds in that period was characterized by an unclear time trend. For England and Sweden, we were unable to find monthly population data so we used yearly population data (Office for National Statistics, 2022; United Nations, 2022) and calculated monthly populations assuming linear growth.

## Estimation framework: Synthetic control method

As patents are enforced and expire nationwide, the gold standard of evaluating health policies using randomized controlled trials (RCTs) cannot be implemented due to the lack of a control group. We therefore draw on a novel quasi-experimental method for causal inference – the synthetic control method (SCM).

SCM is used to assess the impact of an intervention on a treated unit for which a control group is unavailable. Intuitively, SCM creates a synthetic control as a counterfactual by calculating a weighted average of multiple control units. The weights are optimally chosen such that the difference between the treated and the synthetic control unit in the period before the intervention is minimized (Abadie, Diamond and Hainmueller, 2010). The fitting of the weights

in the pre-intervention period is usually based on the pre-intervention outcome values as well as other comparable dimensions (Abadie, 2021).

Control units in the donor pool should fulfill the following three criteria. First, their outcomes should be determined by the same structural process as the treated unit's outcome. Second, they should not be subjected to the same intervention as the treated unit in the study period (Abadie, Diamond and Hainmueller, 2015). Third, there should be no spillover effects from the intervention acting on the treated unit on the control units in the donor pool (Abadie, 2021). We therefore exclude all lipid-modifying agents from the control unit donor pool due to the presence of spillover effects. When the patent of a lipid-modifying agent expired, this typically had a negative impact on the consumption of other lipid-modifying agents. In appendix 3.8 we show the negative impact of atorvastatin's patent expiry on simvastatin consumption in Germany, England and Sweden. Therefore, we do not include any lipid-modifying agents in the control unit donor pool.

We construct our donor pool of control units using molecules acting on the cardiovascular system that are not lipid-lowering agents as they fulfill all three criteria for good donor pool control units. First, these molecules are driven by the same underlying demand for cardiovascular disease prevention as lipid-modifying agents. Second, most molecules did not experience a patent expiry in our observation period (appendix 3.9) and we argue that the few exceptions[29] can be neglected as the weights assigned to individual donor pool molecules are small (appendix 3.14). Lastly, there is no spillover of lipid-modifying agents' patent expiry onto molecules acting on other parts of the cardiovascular system to be expected. Medically, it would not make sense to substitute lipid-modifying agents with antihypertensives.

We arrive at the following donor pools of control molecules: For Sweden, we use all 144 non-lipid-modifying agents acting on the cardiovascular system. For England, cleaning all non-lipid-modifying agents' consumption would have been too time consuming, so that we selected

[29] For example, amlodipine lost its patent protection in 2004 or lercanidipine in 2010 (Klose and Schwabe, 2004; 2010). Appendix 3.9 also shows that some molecules lost their patent protection in 2003 or 2012 but never in the same month as simvastatin or atorvastatin and again, their weights in individual SCM results were small. We could not find the patent expiration dates for all molecules but for the majority.

groups of molecules based on two criteria: First, we chose the molecule groups with the highest consumption to be able to model the high consumption levels of simvastatin and atorvastatin well. Second, the molecule groups selected in the first step had to act as a good synthetic control group for Sweden. In other words, we assumed that a good control group in Sweden would also act as a reasonably good control group in England. We end up with 83 antihypertensives, diuretics, beta-blocking agents and agents acting on the renin-angiotensin system for England. For Germany, cleaning constituted an even higher constraint, so that we followed a similar selection process for its synthetic control group as we did for England, selecting 36 antihypertensives that worked well as synthetic controls in England and Sweden.[30]

We base the weights for constructing the synthetic control groups exclusively on pre-intervention outcome values, the consumed DDD or dispensings per month per 1,000 population. SCM typically employs additional covariates to optimize the synthetic control fit. We do not have other, quantifiable information on the molecules included in our analyses available, so that we fit our synthetic control groups with only the pre-patent expiry molecule consumption. Kaul et al. (2015) discuss the trade-off resulting from this strategy: While we improve the efficiency by using only lagged outcome variables as unobserved confounders are accounted for better, efficiency might be threatened by introducing a small-sample bias. Due to our large number of pre-intervention periods, we believe our estimations to still be sufficiently efficient.

We run our SCM analyses for atorvastatin's patent expiry in England, Germany and Sweden and for simvastatin's patent expiry in Germany. We refer to simvastatin and atorvastatin as the "treated units" with the "intervention" referring to their patent expiry.

[30] In the ATC language, we use DDD molecule consumption from ATC level 2 groups C02 (antihypertensives) and C08 (calcium channel blockers) in Germany. In England, we use molecules from these groups, too, as well as molecules from groups C03 (diuretics), C07 (beta blocking agents) and C09 (agents acting on the renin-angiotensin sytem). For Sweden, all ATC level 2 groups except for the C10 group (lipid modifying agents) is used, ie groups C01 (cardiac therapy), C04 (peripheral vasodilators) and C05 (vasoprotectives) in addition to those used in England.

More formally, each of our synthetic controls are constructed using $J$ molecules in the donor pool. We index molecules with $j = 1, 2, \dots, J, J+1$ where $j = 1$ is the treated molecule. Each donor pool molecule is assigned a weight $w_j$ which is subject to two constraints:

$$W = \begin{bmatrix} w_2 \\ \cdots \\ w_{J+1} \end{bmatrix} \tag{7}$$

Constraint 1: Weights cannot be negative or larger than 1, i.e. $0 \leq w_j \leq 1$ for $j = 2, \ \dots, J$

Constraint 2: The sum of all weights must equal 1, i.e. $w_2 + \cdots + w_{J+1} = 1$

These two weight constraints imply that SCM does not rely on extrapolation (Abadie, 2021) which constitutes an econometric advantage over the Interrupted Time Series methodology.

Weights $W$ are chosen to minimize

$$\min_{w_2,\dots,w_{J+1}} (Y_1 - Y_0 W)^2 \tag{8}$$

where $Y_1$ is a $(1x1)$ matrix with the treated units' pre-intervention outcome, and $Y_0$ is a $(1xJ)$ matrix with the donor pool units' pre-intervention outcome. We denote the optimal weights with $w_j^*$ and index the $T$ number of time periods with $t = 1, 2, \dots T_0, \dots, T$ where $T_0$ is the number of time periods pre-intervention.

In a setting without a true control group like in a randomized controlled trial, one has to rely on the continuity assumption, ie that "no other interventions or confounding covariates [other] than the treatment of interest in analyses changed at the threshold" (Bärnighausen et al., 2017). We therefore examine whether other policy changes such as statin prescription guidelines or health insurances' reimbursement policies occurred at the same time as the patent expiry. We list all relevant guidelines from major US and EU medical associations as well as relevant national associations (and whether they specifically recommend simvastatin or atorvastatin) in appendix 3.10. With three exceptions, none of the relevant guidelines on the prevention and treatment of dyslipidemia and cardiovascular disease recommend specific

lipid-lowering molecules. They usually recommend certain target LDL cholesterol levels – either absolute or percentage reductions - depending on patients' risk factors and CVD history. The British guidelines by the National Institute for Health and Care Excellence (NICE) from 2008 and 2014 and the Swedish guideline by National Board of Health and Welfare (2008) (socialstyrelsen) from 2008 form the exceptions, which specifically recommend simvastatin in the 2008 guidelines and atorvastatin in 2014. Since the timing of their publication does not coincide with simvastatin or atorvastatin's patent expiry and they are non-binding (McGuire and Bauhoff, 2011), we argue that these exceptions also do not threat our continuity assumption. We also presented the three countries' reimbursement systems in the background section and did not find any policy changes to have occurred around the time of the two molecules' patent expiries. We conclude that changes in prescription guidelines or reimbursement policies are unlikely confounders of drug consumption around the time of patent expiry and thereby do not threaten our identification strategies through violations of the continuity assumption.[31]

*Inference*

We use two techniques to perform inference.

First, we follow the permutation inference as introduced by Abadie, Diamond and Hainmueller (2010) by assigning each molecule in the control donor pool the patent expiry treatment. We estimate the synthetic control for the treated molecule and for each donor pool molecule. For each of these iterations, we calculate the Root Mean Squared Prediction Error (RMSPE) for the pre-treatment period:

[31] Additionally, the continuity assumption is more plausibly met in monthly data as the coincidental co-occurrence with other relevant health policy changes is less likely in a period of one month than a year. For this reason, we consider our yearly data only for exploratory, visual inspection and do not proceed to analyze them econometrically. Additionally, Zhang, Wagner and Ross-Degnan (2011) showed that interrupted time series studies with few time points, among other study characteristics, may be underpowered and caution against the interpretation of such results.

$$RMSPE_{pre} = \left( \frac{1}{T_0} \sum_{t=1}^{T_0} \left( Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt} \right)^2 \right)^{\frac{1}{2}} \tag{9}$$

We compute the RMSPE for the post-treatment period, as well, and then calculate the ratio of the post- to pre-treatment RMSPE. Finally, we sort all ratios in descending order, generate a ranking and use the rank of the treated molecule to calculate the $p$-value:

$$p = \frac{RANK}{TOTAL} \tag{10}$$

However, as Abadie points out "even if a synthetic control is able to closely fit the trajectory of the outcome variable for the treated unit before the intervention, the same may not be true for all the units in the donor pool" (2021). We follow Firpo and Possebom (2018) and discard molecules from the permutation distribution that have pre-intervention MSPEs[32] that are more than five times larger than that of the treated unit. This is because placebo studies for those molecules would not be informative for assessing whether the post-patent consumption of atorvastatin could have occurred by chance. We consider $p$-values lower than 0.05 statistically significant.

Second, we conduct a falsification exercise that Abadie, Diamond and Hainmueller (2015) refer to as "in-time placebo" where we backdate patent expiries to an earlier date. According to Opartny (2021), one can arbitrarily backdate the intervention, so we backdate the patent expiry by two years. In other words, we pretend that the patent expiry happened two years before its actual date. Our results would then be considered robust if (i) the synthetic control still closely tracks the treated unit until the intervention, and (ii) a gap still emerges after the intervention (Abadie and Vivies-i-Bastida, 2022).

## Robustness checks

We corroborate our results with several robustness checks. First, we run two sensitivity analyses of our synthetic control method results by way of two donor pool modifications.

[32] Mean Squared Prediction Errors (MSPEs) are squared RMSPEs.

Second, we replicate our results with the method that other studies examining the impact of patent expiry on drug consumption have used, namely interrupted time series (ITS).

*Sensitivity analysis of the synthetic control analysis*

We first ensure that our results do not rely on the inclusion of one individual control molecule. We do so by ensuring that no donor pool molecule was assigned a disproportionately large weight. We then also calculate leave-one-out estimates of our main SCM models by iteratively dropping one donor pool control molecule as introduced by Abadie, Diamond and Hainmueller (2015). The leave-one-out estimates should all mirror the original results based on the entire donor pool.

Second, we reduce our donor pool to eight molecules. We use the most dispensed molecule within each molecule group for Sweden, the two most consumed molecules within each molecule group for England, and the four most consumed molecules within each molecule group for Germany. This is to address the potential presence of an over-fitting or an interpolation bias due to the large number of control donor pool units in our main analysis (Abadie and Vivies-i-Bastida, 2021). We use the most consumed molecules as the lipid-modifying molecules under investigation are characterized by large consumption volumes so that the donor pool molecules with the highest consumption will be the molecules most similar to simvastatin and atorvastatin.

*Replication with interrupted time series*

An interrupted time series (ITS) design examines changes in a time series in response to a real-world event or policy change (Wagner et al., 2002; Bernal, Cummins and Gasparrini, 2017). ITS has been widely used to evaluate public health interventions ranging from bicycle helmet laws and new vaccine introductions to the prevention of bacterial infections in intensive care units (Dennis et al., 2012; Lau et al., 2015; Derde et al., 2013) as well as the impact of patent expiry on drug consumption (Chapman, Fitzpatrick and Aladul, 2017; Fiorentini, Bruni and Mammi, 2022).

We apply the interrupted time series design in its simplest form, using the following regression:

$$Y_t = \beta_0 + \beta_1 T_t + \beta_2 X_t + \beta_3 Z_t \quad (11)$$

Where $Y_t$ is the population-corrected DDD per month consumption of the molecule of interest, $T_t$ represents the number of months elapsed since the start of the observation period, $X_t$ is a dummy variable for the time periods before (coded 0) and after (coded 1) the patent expiry, and $Z_t$ represents the number of months since the patent expiry, taking value 0 before it. The coefficients $\beta_0$, $\beta_1$, $\beta_2$ and $\beta_3$ represent the baseline molecule consumption at $T_0$, the time trend pre-patent expiry, and the level ('jump') and trend ('slope') changes in response to the patent expiry, respectively. Our a-priori impact model is to expect a trend but not necessarily a level change. This is because of inertia in prescription and drug consumption habits (Coscelli, 2000). In other words, one is likely to observe switches in the molecules prescribed to new patients initiating a statin therapy for the first time while patients already on a statin regiment are unlikely to switch. The previous literature also reports slope but no stark level changes in statin consumption in response to patent expiries (Chapman, Fitzpatrick and Aladul, 2017; Fiorentini, Bruni and Mammi, 2022).

We conduct ITS analyses for a period of 80 months (± 40 months around expiry) and, for robustness, also for 60 months (± 30 months around expiry). We use symmetric time frames around the patent expiry because Zhang, Wagner and Ross-Degnan (2011) use simulations to show that these maximize statistical power.

We use extrapolation to predict counterfactual values as if patent expiry had not taken place. After estimating equation (1), we use its fitted intercepts $\hat{\beta}_0$ and $\hat{\beta}_1$ to estimate the counterfactual molecule consumption $Y_t^{counter}$ for all $t$ following the patent expiry as follows:

$$Y_t^{counter} = \hat{\beta}_0 + \hat{\beta}_1 T_t \quad (12)$$

Seasonality and autocorrelation are two major issues that should be considered when conducting ITS analyses (Bhaskaran et al., 2013).

We do not undertake special measures to take care of seasonality for two reasons: First, an initial descriptive analysis as recommended by Bernal, Cummins and Gasparrini (2017) does not provide a strong indication for seasonal patterns in molecule consumption. Intuitively, this makes sense as the indication for a statin prescription - cardiovascular disease risk - does not vary seasonally like other drug indications such as the indication of flu vaccinations - prevalence of flu infections - do. Second, seasonality is most problematic when being interested in short-term outcomes which could be confounded by longer-term seasonal patterns but as this paper is predominantly concerned with examining the longer-term impact of patent expiry on molecule consumption rather than its immediate impact (jumps in consumption or slope increases after few months), we argue that the ITS in this application is unlikely to be threatened by seasonality.

We examine the presence of autocorrelation by visual inspection of the residual plots as well as the autocorrelation function (ACF) and partial autocorrelation function (PACF) graphs and also test for autocorrelation with the Durbin Watson test (Savin and White, 1977; see appendix 3.11). The residual plots and the Durbin Watson test do not strongly reject autocorrelation, but the PACF shows several significant partial autocorrelations, clearly indicating the presence of autocorrelation. Therefore, we correct for autocorrelation using Newey-West corrected standard errors (Newey and West, 1987). We include 3 lags because the number of lags should be approximately equal to $t^{1/4}$ where $t$ corresponds to the number months (t=80) so that $80^{1/4} = 2.991 \approx 3$ (Greene, 2012).

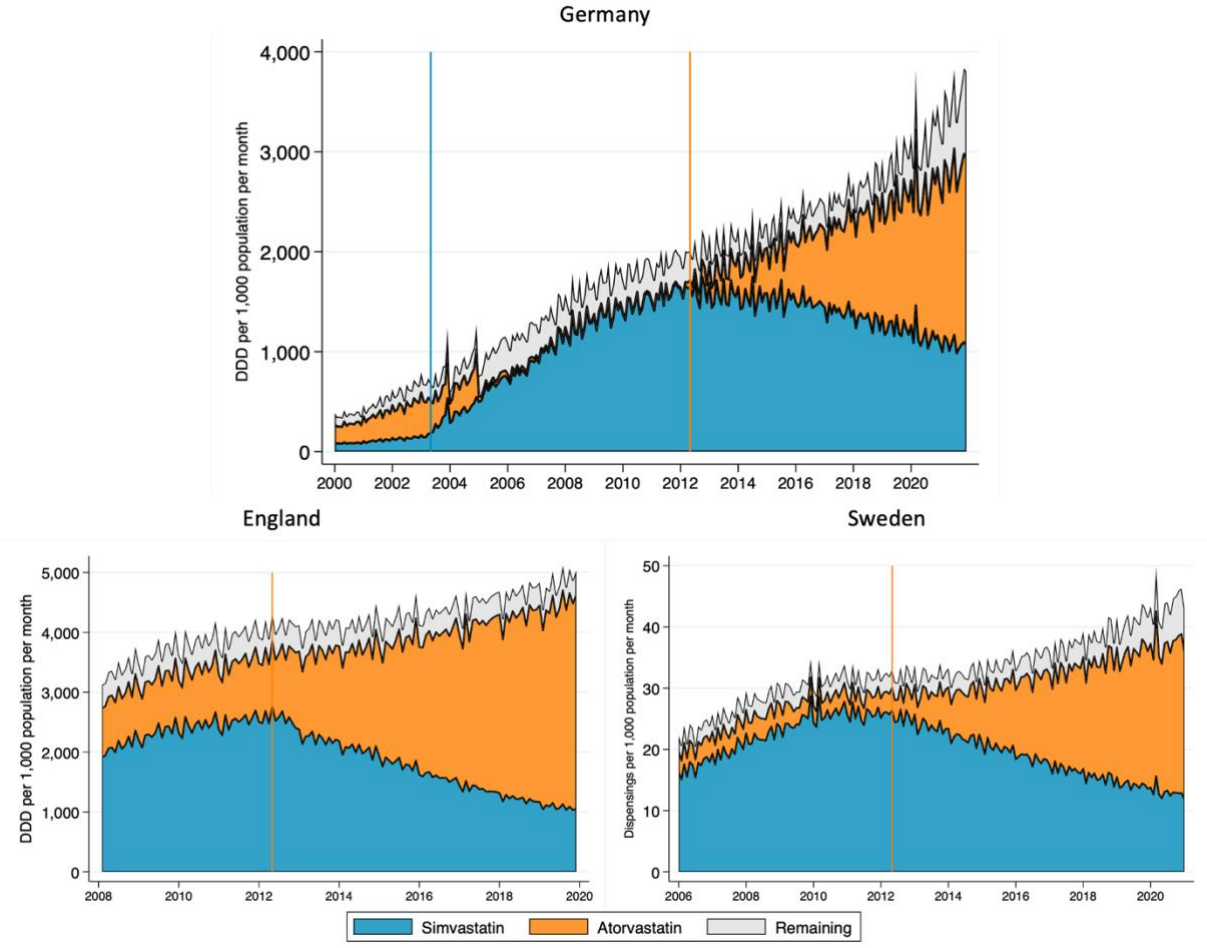
We do not implement an ITS extension that accounts for bias arising from time-varying confounders and is known as controlled interrupted time series (Bernal, Cummins and Gasparrini, 2017) as it is a simplified version of the synthetic control method. It involves the addition of at least one control group with similar level and trend of the dependent variable and pre-intervention covariates and then applies a propensity score-based weighted regression model (Linden and Adams, 2011).

## 4.5. Results

### Overall lipid-modifying agent consumption

Figure 9 presents the lipid-modifying agents' consumption of Germany, England and Sweden which simvastatin and atorvastatin clearly dominate in all three countries.[33] Overall, we see an increase in the consumption of lipid-modifying agents in all three countries. In Germany, total lipid-modifying agent consumption increased from 385 DDD per 1,000 population in January 2000 to 3,802 in December 2021. In England, we see an increase from 3,124 DDD per 1,000 population in February 2008 to 5,033 in December 2019. The number of per 1,000 population dispensings in Sweden increased from 22 dispensings in January 2006 to 43 in January 2021.

*Figure 9: Monthly lipid-modifying agent consumption in Germany, England and Sweden*




Note: Vertical lines refer to patent expiry of molecule in same color.

[33] Appendix 3.12 shows the statin consumption for all countries that we could find only yearly data for.

In Germany, we see that atorvastatin was initially responsible for a higher consumption share than simvastatin before March 2003. This is when the first lipid-modifying agent, simvastatin, lost its patent protection. From March 2003 onwards, simvastatin's market share increased substantially and that of atorvastatin started decreasing and dropped to almost zero in January 2005. Before atorvastatin's patent expiry in May 2012, simvastatin was the predominantly consumed molecule in all three countries. With its patent expiry, the trend reversed, with atorvastatin gradually displacing simvastatin in all three countries.

## Impact of patent expiry on individual lipid-modifying agents' consumption

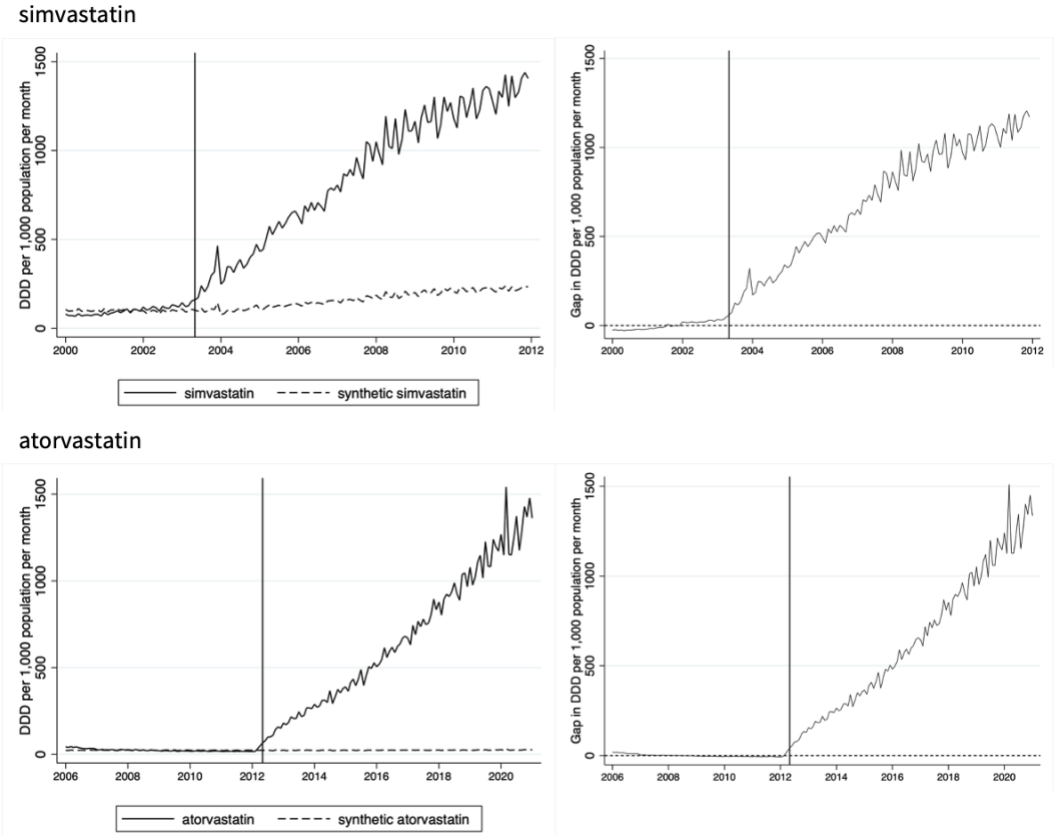
We display our synthetic control results in Figure 10 for Germany and Figure 11 for England and Sweden.[34] We list the weights assigned to each donor pool molecule in the appendix 3.13. The figures' right panels additionally display the gap graphs which plot the difference between the actual and synthetic consumption.

Two major findings emerge:
First, the synthetic control method is an appropriate tool for examining the impact of patent expiry on molecule consumption as we find that the synthetic control closely fits the actual molecule consumption prior to the molecule's patent expiry, for all countries and molecules. This close fit is especially remarkable as we only use the pre-intervention outcome variables to fit the weights for the synthetic control group and do not use additional covariates like other SCM papers usually do (Abadie, 2021). This is an important methodological improvement for the literature on the impact of patent expiry on drug consumption as the interrupted time series method relies on linear trends and extrapolation whereas the synthetic control method does not.

Second, the SCM results confirm the stark increase in molecule consumption post patent expiry. A year after the patent expiry, simvastatin and atorvastatin consumption increased by

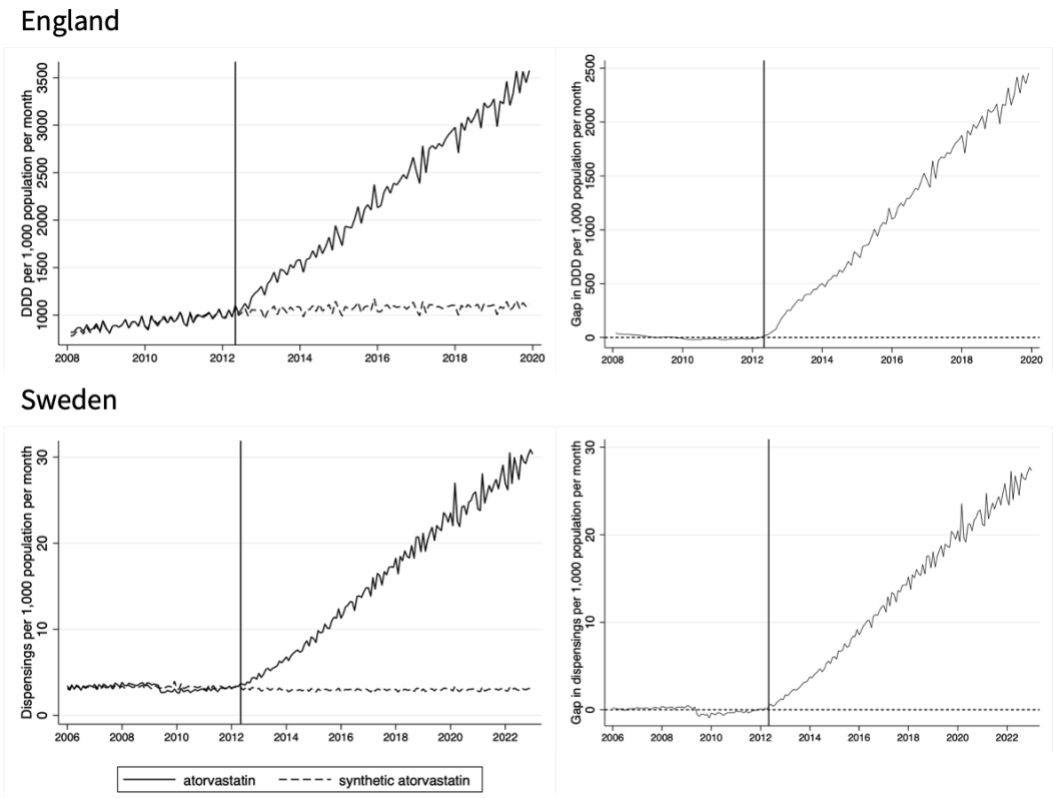
[34] Line graphs of individual monthly simvastatin and atorvastatin consumption are included in appendix 3.13.

*Figure 10: Consumption of simvastatin and atorvastatin versus their synthetic control in Germany*




Note: Vertical lines correspond to molecule's patent expiry. Left panel: molecule vs. synthetic molecule; right panel: consumption gap between molecule and synthetic molecule

*Figure 11: Consumption of atorvastatin versus their synthetic control in England and Sweden*




Note: Vertical line indicates date of atorvastatin's patent expiry. Left panel: atorvastatin vs. synthetic control; right panel: consumption gap between atorvastatin and synthetic control

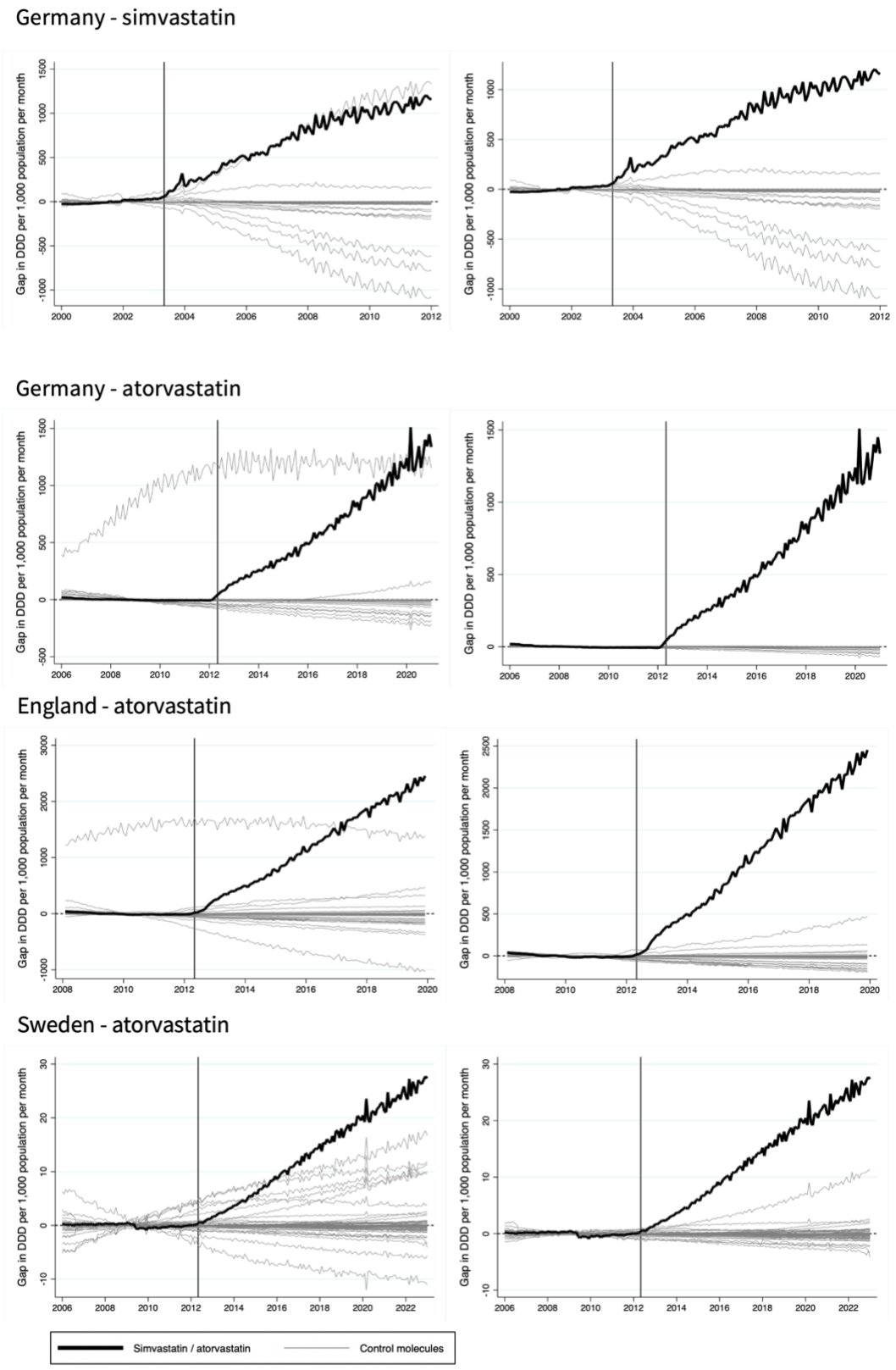
241% and 788% in Germany, respectively. After five years, consumption increased by 475% and 2,955%, respectively. In England, atorvastatin consumption one year after patent expiry was 53% higher and 148% higher five years later. In Sweden, atorvastatin consumption one year after patent expiry was 75% higher and 536% higher five years later.

Our inference methods find our results to be statistically significant.
Our permutation distributions for calculating exact p-values from these Placebo estimates are shown in Figure 12. Compared against the consumption distribution of control molecules, the gap for simvastatin in Germany and atorvastatin in England, Germany and Sweden appears highly unusual, especially when considering the right panel which drops control group molecules whose pre-patent expiry fit was bad.[35] The distribution of post/pre-patent expiry MSPEs confirms this impression, as shown in Figure 13, where the ratio of simvastatin and atorvastatin clearly stand out and all four country-molecule combinations under consideration yield p-values smaller than 0.05 which we consider to be statistically significant.

Our predating of the intervention likewise corroborates our main results as the synthetic control still (i) closely tracks simvastatin or atorvastatin's consumption up to and (ii) diverges into an increasing gap after the true patent expiry date (see figure 14).

[35] In other words, we discard all placebo distributions of control group molecules where their pre-patent expiry MSPE ratio was five times higher than that of simvastatin / atorvastatin.

*Figure 12: Consumption gaps in simvastatin / atorvastatin versus placebo gaps in control molecules*




Notes: Vertical line indicates date of molecule's patent expiry. Left panel: Atorvastatin / simvastatin and all control molecules; right panel: Atorvastatin / simvastatin and control molecules after discarding pre-patent expiry MSPE five times higher than that of molecule.

*Figure 13: Ratio of post- to pre-patent expiry RMSPEs of simvastatin / atorvastatin vs. donor pool control molecules*

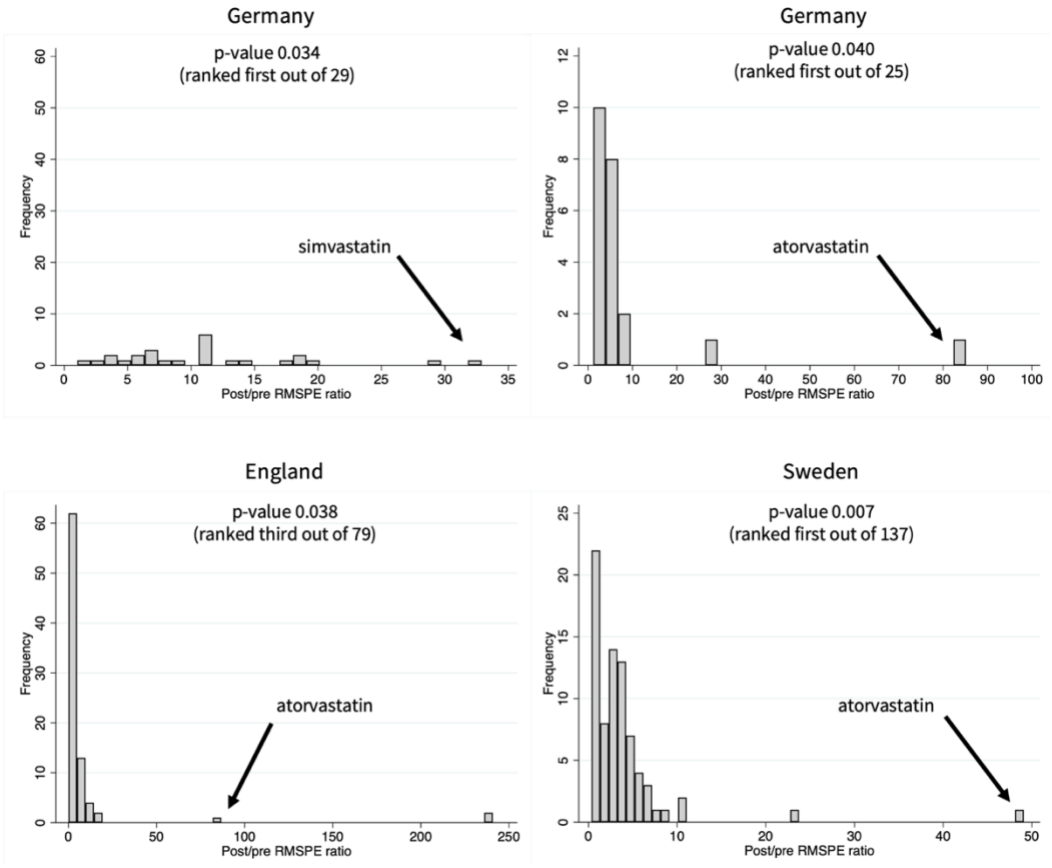
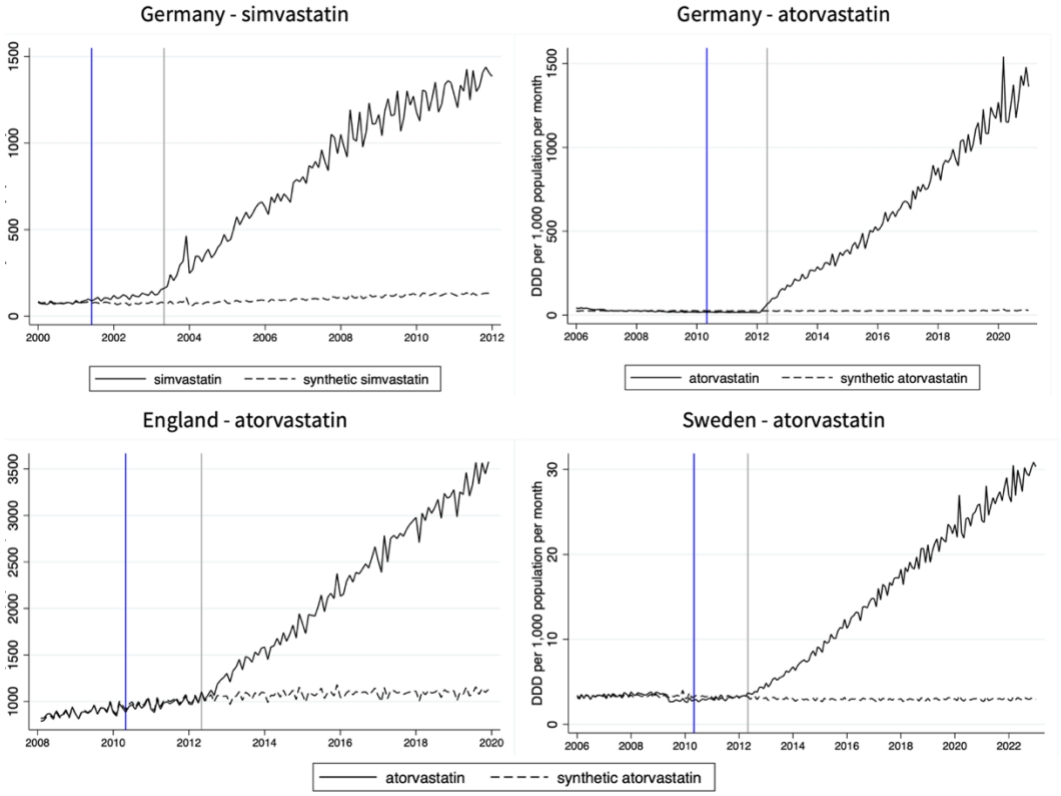



*Figure 14: Predating the patent expiry of simvastatin to May 2001 and of atorvastatin to May 2010.*




Note: Blue vertical lines correspond to Placebo patent expiry date and grey vertical lines correspond to true patent expiry dates.

## Discussion of mechanisms

The increase in overall lipid-modifying agent consumption that we see in all three countries is probably driven by a lowering of the thresholds of who should receive preventative statin treatment rather than an increase in the dosages per patient. While the number of individuals on lipid-modifying agent regiments could be consuming increasingly higher dosages, we simultaneously see the number of prescriptions, a measure independent of the dosage of a drug, to increase in the countries where we have prescription data, England and Sweden. We therefore argue that increasing dosages probably do not explain the increasing lipid-modifying agent consumption by itself. A more promising explanation for the observed increase would be that the number of individuals consuming lipid-modifying agents preventatively has increased. While we do not have access to clinical data that would confirm this, we argue that this is a likely explanation as medical guidelines have continuously lowered the thresholds on who to test for CVD risk and who to prescribe a statin treatment to.[36]

Another explanation for an increasing consumption of lipid-modifying agents could be an increasing disease burden. However, the relationship between CVD disease burden and lipid-modifying agent consumption is causal in both directions, which renders a visual inspection of the number of DALYs associated with CVD in Germany, England or Sweden insufficient to determine the role of the disease burden in the increasing lipid-modifying agent consumption. In other words, the stagnant CVD disease burden could be observed because of or despite of an increasing consumption of lipid-modifying agents.

The increase in the consumption of individual lipid-modifying agents upon their patent expiry and the concurrent displacement of other lipid-modifying agents is predominantly driven by prices, with reimbursement guideline changes playing a reinforcing role. In Germany, we see that simvastatin or atorvastatin's shares of consumption are negatively associated with their respective prices. The predominant consumption of atorvastatin before 2003 can be explained by branded atorvastatin being available more cheaply. Once generic simvastatin became

[36] For example, the British NICE guidelines from 2008 recommended a lipid-modifying agent treatment for everyone having a calculated risk of CVD of at least 20% over ten years (NICE, 2008). In 2014, the British NICE guidelines already recommended a lipid-modifying agent treatment for all those with a calculated CVD risk of at least 10% over ten years.

available from March 2003 onwards[37], simvastatin was available for a much lower price than generic atorvastatin and started being consumed much more.[38] When the patent of atorvastatin expired in May 2012[39], however, generic atorvastatin became available for an even lower price per DDD than generic simvastatin, explaining why atorvastatin started displacing simvastatin. We observed similar price drops of atorvastatin in England and Sweden upon its patent expiry. These results indicate that all three health systems under consideration are sensitive to price changes and practice substitution when a cheaper lipid-modifying agent is available, ie they exhibit a high elasticity of substitution.

The probability of side effects or the ease of use, on the other hand, do not seem to be associated with individual lipid-modifying agents' consumption.

All three health systems' high elasticity of substitution between simvastatin and atorvastatin means that they neglect the differences in the probability of developing rhabdomyolysis, severe muscle pain, as a side effect or in the ease of use (simvastatin has to reliably be taken at night while atorvastatin does not, see section 4.3). In other words, even though atorvastatin, a lipid-modifying agent with less likely side effects and easier usage, was available before its expiry in May 2012 already, health systems preferred the prescription and reimbursement of simvastatin. Some medical guidelines mentioned these differences but did not proceed to actively recommend atorvastatin vis-à-vis simvastatin (eg Stone et al., 2013).

We see a significant impact of changes to the reimbursement system in January 2005 in Germany. This is when a reference price, ie a single reimbursement amount, was introduced for all lipid-modifying agents so that patients suddenly faced much higher co-payment costs

[37] Two generic simvastatin products became available in March 2003 already as the simvastatin patent-holder Merck, introduced its own generic, Zocor MSD, and also sold its early-entry rights to a generics producer (McGuire and Bauhoff, 2011). By the end of 2003, 29 different simvastatin products were available already, generating an estimated €220 million in cost savings (Klose and Schwabe, 2004b).

[38] Lovastatin, the first statin that was available on the German market in 1989, never played a meaningful role in the German statin market as it became generically available only in 2004 and for higher prices than simvastatin: 1.89€/DDD in 2002 pre-patent expiry and 0.54€/DDD for generics post-patent expiry in 2004 (Klose and Schwabe, 2003; 2006).

[39] Pfizer granted two pharmaceutical firms, Hexal AG and Basics GmbH, a daughter of the Indian company Ranbaxy, early entry rights from March 2012 onwards (Pharmazeutische Zeitung, 2012), which explains the rise in atorvastatin consumption a few months before the official patent expiry date in May 2012.

for atorvastatin than for simvastatin (Dylst, Vulto and Simoens, 2011). The change in reimbursement policy therefore played a reinforcing role of prices in determining Germany's lipid-modifying agent consumption. Atorvastatin was already being displaced after simvastatin's patent expiry in May 2003 but its consumption only dropped to almost zero in January 2005.

Overall, the observed increased molecule consumption post patent expiry in Germany, England and Sweden is driven by price decreases due to the entry of generic producers and was reinforced by a reimbursement policy change in Germany in January 2005.

## Robustness

We first discuss the replication of our results with a different unit in England and the implications thereof for the robustness of our results from Sweden. Then, we present two robustness exercises of our SCM results and our complementary analyses using ITS. Across all robustness check specifications, we find the results to be very similar to our main results.

### *Replication of DDD results using prescription data in England*

We find the composition of the total lipid-modifying agent consumption in England to be similar when using the number of DDD consumed or the number of prescriptions filed per month per 1,000 population (appendix 3.15). We also replicate the individual molecules' consumption increases with ITS analyses. This has implications for the use of the number of dispensings in our Swedish data which we could not convert to the number of consumed DDD. With the robustness of our English results to using the number of prescriptions instead of the number of DDD, we argue that the Swedish results are likely to have been robust to using the consumed DDD as the unit of analysis as well and that our results are unlikely to be threatened by the Swedish data employing a different unit to the German or English data.

### *Robustness exercises for the synthetic control method results*

Our SCM results on the impact of patent expiry on individual molecule consumption are robust to removing individual molecules from the donor pool, i.e. the good pre-patent expiry

fit does not rely on the inclusion of specific individual molecules in the donor pool (see appendix 3.16). All leave-one-out synthetic controls are very similar to the original synthetic controls and thereby yield similarly large molecule consumption increases as in the original analysis.

The results likewise pass the robustness check of reducing our donor pool to just eight molecules (see appendix 3.17). With the smaller donor pools, we still see a good pre-patent expiry fit of the synthetic control unit to the actual molecule consumption and again find molecule consumption to increase substantially post-patent expiry. We therefore argue that our results are not substantially threatened by overfitting or interpolation bias.

*Interrupted Time Series*

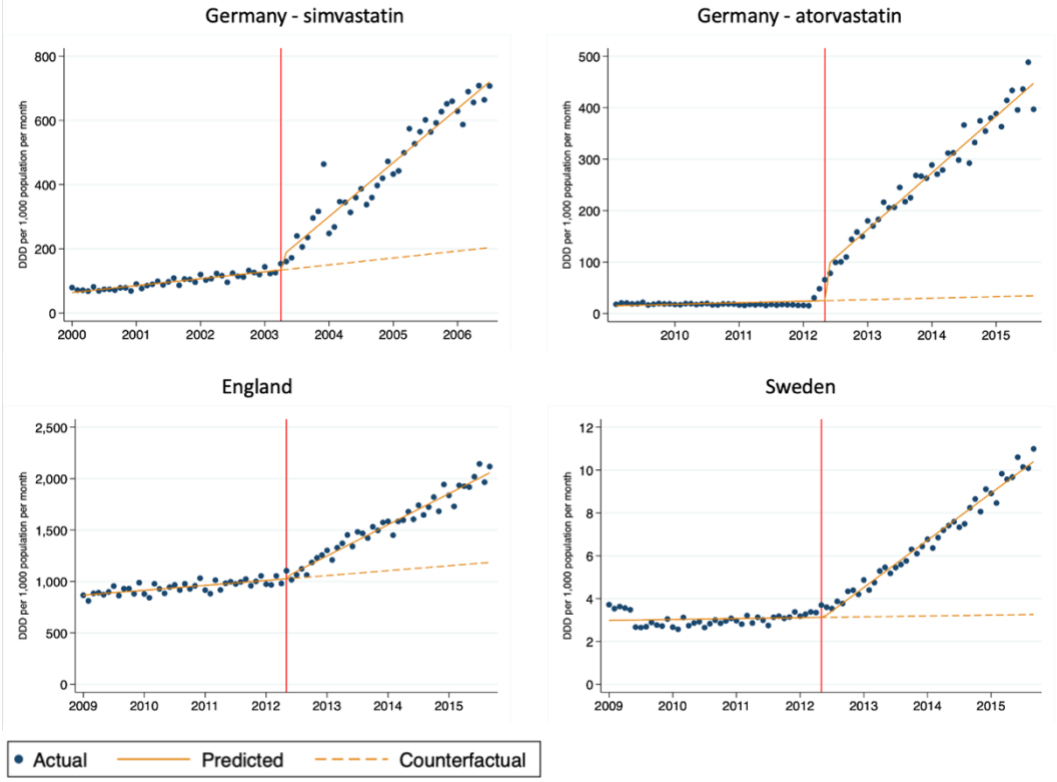
We present the interrupted time series results for simvastatin consumption in Germany and atorvastatin consumption in Germany, England and Sweden in Figure 15 and table 10. For all country-molecule combinations, we see a statistically and economically large increase in the monthly consumption trend post patent expiry. Even though there are also statistically significant jumps in Germany's simvastatin and atorvastatin consumption at their respective patent expiries, these are not as economically meaningful because they are the equivalent of only six and seven months of the time trend increase.

*Table 10: Interrupted time series results of the impact of simvastatin and atorvastatin's patent expiry on their consumption*

| | simvastatin Germany (1) | atorvastatin Germany (2) | atorvastatin England (3) | atorvastatin Sweden (4) |
|---|---|---|---|---|
| time ($\beta_1$) | 2.26*** | 0.294 | 3.97*** | 0.006 |
| | (0.176) | (0.290) | (0.429) | (0.008) |
| expiry ($\beta_2$) | 69.4*** | 77.6*** | 21.4 | -0.107 |
| | (20.8) | (12.3) | (22.7) | (0.156) |
| time post expiry ($\beta_3$) | 13.1*** | 10.3*** | 21.3*** | 0.177*** |
| | (0.807) | (0.346) | (0.868) | (0.010) |
| constant ($\beta_0$) | 69.6*** | 17.5*** | 867*** | 2.91*** |
| | (3.48) | (4.15) | (10.1) | (0.214) |
| Observations | 80 | 80 | 80 | 80 |

Note: We employ Newey-West standard errors correcting for heteroscedasticity and autocorrelation and display statistical significance levels (10%, 5%, 1%) with */**/***.

*Figure 15: The impact of simvastatin and atorvastatin's patent expiry on their respective consumption using interrupted time series analysis*




Note: The vertical red lines correspond to the date of the molecule's patent expiry.

We include the same analyses with using a smaller, 60-month time period in appendix 3.18.

## 4.6. Conclusion

This paper has examined the impact of patent expiry on the German, English and Swedish consumption of simvastatin and atorvastatin, the two most consumed lipid-modifying agents in the three countries. Lipid-modifying agents are prescribed to prevent the development of cardiovascular disease. They lower the cholesterol levels in the blood, thereby prevent the build-up of plaque in the arteries which could otherwise result in cardiovascular disease events like strokes or heart attacks. The European patents for simvastatin and atorvastatin expired in May 2003 and May 2012, respectively.

Other papers examining the impact of patent expiry have typically focused on the impact on prices (review by Vondeling et al., 2018) or on the market share of the branded vis-à-vis the

generic products (see footnote 14). The small literature that studies the impact on the overall consumption of a molecule whose patent expired has been US-centric, found mixed effects and employed descriptive, regression or (interrupted) time-series analyses (Aitken, Berndt and Cutler, 2009; Aitken et al., 2013; Berndt, Kyle and Ling, 2003; Chapman, Fitzpatrick and Aladul, 2017; Duflos and Lichtenberg, 2012; Fiorentini, Bruni and Mammi, 2022; Imai, Fushimi and Sundell, 2018; Lakdawalla and Philipson, 2012). Here, we present evidence from Germany, England and Sweden employing the synthetic control method, a quasi-experimental tool which has never been used in the literature on the impacts of patent expiry before. The excellent fit of our synthetic controls prior to the patent expiry of simvastatin and atorvastatin indicates that this research question is well-suited to the utilization of the synthetic control method. We therefore contribute novel empirical evidence as well as a methodological advancement to this literature.

We find that the consumption of simvastatin and atorvastatin increases upon their patents' expiration in all three countries. We simultaneously observe a displacing effect on other lipid-modifying agents. We argue that the main determinant of this substitution is the price decrease effected by the entry of generic producers upon patent expiry and that changes in reimbursement policies can reinforce this effect of prices. Countries' price sensitivity and thereby high elasticity of substitution between simvastatin and atorvastatin led to substantial cost savings but simultaneously neglected the increased likelihood of a severe side effect with the predominant consumption of simvastatin between 2003 and 2012.

Future research should extend this literature in multiple ways.
First, the results from this study should be replicated with evidence from more countries and other drug classes, preferably using high frequency such as monthly data.
Second, the roles of specific health system characteristics should be examined by building an evidence base from multiple countries. We showed that a specific reimbursement policy change had a reinforcing effect on the impact of patent expiry on drug consumption in January 2005 in Germany. It is important to understand the role of other features like co-payment rules, cost-saving tools like GP-level drug quotas or drugs' accessibility over the counter vis-à-vis prescriptions in determining the impact of patent expirations on drug consumption.

Third, the impact that an individual molecule's patent expiry has on the consumption of its whole drug class should also be examined.[40] We expect this impact to be relatively small in countries with well-functioning social health insurance systems such as Germany, England or Sweden. In privatized healthcare systems like the United States or in low- and middle-income countries with an overall lower health insurance coverage, on the other hand, we would expect the impact on drug class consumption to be larger. In other words, patent expiry in these contexts would not only result in substitution at the molecule level within a drug class but probably also result in an extensive margin increase in drug consumption.[41]

This would have significant implications for the health outcomes that the drugs whose patent expires target. Examining the impact of patent expiry on health outcomes is therefore the fourth and last future research avenue we would like to recommend. Policy makers should understand the health implications of monopoly drug prices for the duration of the patent protection – typically twenty years – when debating how to incentivize pharmaceutical innovation while also providing the best medical care at an affordable cost, be it with a patent system or otherwise (Ahmad, Naeher and Vollmer, 2023).

[40] To conduct this research with a SCM design would require a different donor pool as the consumption of a whole drug class is much larger than that of individual molecules so that a weighted average of individual molecules would not result in a well-fitting synthetic control. Instead, the donor pool would have to be made up of the consumption of other drug *classes*.

[41] Descriptively, this is what we see when considering the impact of simvastatin's patent expiry in the United States on 23 June 2006, the first lipid-modifying agent to lose patent protection, where total lipid-modifying agent consumption increased dramatically from 2007 onwards (see appendix 3.11).

# 5. References


Abadie, A., 2021. Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, *59*(2), pp.391-425.

Abadie, A., Diamond, A., and Hainmueller, J., 2010. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, *105*(490), pp. 493-505.

Abadie, A., Diamond, A., and Hainmueller, J., 2015. Comparative politics and the synthetic control method. *American Journal of Political Science*, *59*(2), pp. 495-510.

Abadie, A., and Gardeazabal, J., 2003. The economic costs of conflict: A case study of the Basque Country. *American economic review*, *93*(1), pp. 113-132.

Abadie, A., and Vives-i-Bastida, J., 2022. Synthetic controls in action. *arXiv preprint arXiv:2203.06279*.

Ahmad, A., Naeher, D. and Vollmer, S., 2023. The Political Economy of Patent Buyouts. Courant Research Centre 'Poverty, Equity and Growth' Discussion Paper No. 290.

Aitken, M., Berndt, E.R. and Cutler, D.M., 2009. Prescription Drug Spending Trends In The United States: Looking Beyond The Turning Point: The drug spending trends observed in the 1980s, 1990s, and the first few years of this decade have changed dramatically in the past five years—bringing both opportunity and threat. *Health Affairs*, *27*(Suppl1), pp.w151-w160.

Aitken, M.L., Berndt, E.R., Bosworth, B., Cockburn, I.M., Frank, R., Kleinrock, M. and Shapiro, B.T., 2013. The regulation of prescription drug competition and market responses: patterns in prices and sales following loss of exclusivity. In *Measuring and Modeling Health Care Costs* (pp. 243-271). University of Chicago Press.

Almond, D., Mazumder, B. and Van Ewijk, R., 2015. In utero Ramadan exposure and children's academic performance. *The Economic Journal*, *125*(589), pp.1501-1533.

Angrist, N., Evans, D.K., Filmer, D., Glennerster, R., Rogers, F.H. and Sabarwal, S., 2020. *How to Improve Education Outcomes Most Efficiently?: A Comparison of 150 Interventions Using the New Learning-adjusted Years of Schooling Metric*. World Bank Group, Education Global Practice & Development Research Group.

Anderson, M.L., 2008. Multiple inference and gender differences in the effects of early intervention: a reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association, 103*(484), pp. 1481–1495.

Angrist, J.D. and Lavy, V., 1999. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly journal of economics*, *114*(2), pp.533-575.

Angrist, N., Evans, D. K., Filmer, D., Glennerster, R., Rogers, F. H., and Sabarwal, S. (2020). How to improve education outcomes most efficiently? A Comparison of 150 interventions using the new Learning-Adjusted Years of Schooling metric. *World Bank Policy Research Working Paper,* (9450).

Angrist, N. and Meager, R., 2023. Implementation Matters: Generalizing Treatment Effects in Education. *EdWorkingPaper No. 23-802.*

Arnett, D. K., Blumenthal, R.S., Albert, M.A., Buroker, A.B., Goldberger, Z.D., Hahn, E.J. Himmelfarb, C.D., Khera, A., et al., 2019. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease. *Circulation,* 140(11), pp. e596-e646.

Ashwell, M., Gunn, P. and Gibson, S., 2012. Waist-to-height ratio is a better screening tool than waist circumference and BMI for adult cardiometabolic risk factors: systematic review and meta-analysis. *Obesity reviews*, *13*(3), pp. 275-286.

Bae, J.P., 1997. Drug patent expirations and the speed of generic entry. *Health services research*, *32*(1), p.87.

Bärnighausen, T., Oldenburg, C., Tugwell, P., Bommer, C., Ebert, C., Barreto, M., Djimeu, E., Haber, N., et al., 2017. Quasi-experimental study designs series—paper 7: assessing the assumptions. *Journal of clinical epidemiology*, *89*, pp.53-66.

Baird, S., Hicks, J.H., Kremer, M. and Miguel, E., 2016. Worms at work: Long-run impacts of a child health investment. *The Quarterly Journal of Economics*, *131*(4), pp.1637-1680.

Baird, S., McIntosh, C. and Özler, B., 2011. Cash or condition? Evidence from a cash transfer experiment. *The Quarterly journal of economics*, *126*(4), pp.1709-1753.

Banerjee, A.V., Banerji, R., Duflo, E., Glennerster, R. and Khemani, S., 2010. Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India. *American Economic Journal: Economic Policy*, *2*(1), pp.1-30.

Banerjee, A. V., Cole, S., Duflo, E. and Linden, L., 2007. Remedying education: Evidence from two randomized experiments in India. *The Quarterly Journal of Economics*, *122*(3), pp. 1235-1264.

Barrera-Osorio, F., Fasih, T. , Patrinos, H. A. and Santibáñez, L., 2009. *Decentralized decision-making in schools: The theory and evidence on school-based management*. Washington DC: The World Bank.

Beasley, E. and Huillery, E., 2017. *Willing but unable? short-term experimental evidence on parent empowerment and school quality*. The World Bank.

Bauhoff, S., 2014. The effect of school district nutrition policies on dietary intake and overweight: a synthetic control approach. *Economics & Human Biology*, *12*, pp.45-55.

Benhassine, N., Devoto, F., Duflo, E., Dupas, P. and Pouliquen, V., 2015. Turning a shove into a nudge? A "labeled cash transfer" for education. *American Economic Journal: Economic Policy*, *7*(3), pp.86-125.

Benjamini, Y. and Hochberg, Y., 1995. *Controlling the false discovery rate: a practical and powerful approach to multiple testing*. Journal of the Royal Statistical Society: Series B (Methodological), 57 (1), 289–300.

Bernal, J.L., Cummins, S. and Gasparrini, A., 2017. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *International journal of epidemiology*, *46*(1), pp.348-355.

Berndt, E.R., Kyle, M. and Ling, D. 2003. The Long Shadow of Patent Expiration. Generic Entry and Rx-to-OTC Switches. In *Scanner Data and Price Indexes* (pp. 229-273). University of Chicago Press.

Beuermann, D.W., Cristia, J., Cueto, S., Malamud, O. and Cruz-Aguayo, Y., 2015. One laptop per child at home: Short-term impacts from a randomized experiment in Peru. *American Economic Journal: Applied Economics*, *7*(2), pp.53-80.

Bhaskaran, K., Gasparrini, A., Hajat, S., Smeeth, L. and Armstrong, B., 2013. Time series regression studies in environmental epidemiology. *International journal of epidemiology*, *42*(4), pp.1187-1195.

Blais, J.E., Wei, Y., Yap, K.K., Alwafi, H., Ma, T.T., Brauer, R., Lau, W.C., Man, et al., 2021. Trends in lipid-modifying agent use in 83 countries. *Atherosclerosis*, *328*, pp.44-51.

Bleakley, H., 2010. Malaria eradication in the Americas: A retrospective analysis of childhood exposure. *American Economic Journal: Applied Economics*, *2*(2), pp.1-45.

Blimpo, M.P., Evans, D. and Lahire, N., 2015. Parental human capital and effective school management: evidence from The Gambia. *World Bank Policy Research Working Paper*, (7238).

Böhm, K., Schmid, A., Götze, R., Landwehr, C. and Rothgang, H., 2013. Five types of OECD healthcare systems: empirical results of a deductive classification. *Health policy*, *113*(3), pp.258-269.

Boersma, C., Klok, R.M., Bos, J.M., Naunton, M., Van Den Berg, P.B., de Jong-van den Berg, L.T. and Postma, M.J., 2005. Drug costs developments after patent expiry of Enalapril, Fluoxetine and Ranitidine: a study conducted for the Netherlands. *Applied health economics and health policy*, *4*, pp.191-196.

Bold, T., Filmer, D., Martin, G., Molina, E., Stacy, B., Rockmore, C., Svensson, J. and Wane, W., 2017. Enrollment without learning: Teacher effort, knowledge, and skill in primary schools in Africa. *Journal of Economic Perspectives*, *31*(4), pp.185-204.

Bold, T., Kimenyi, M., Mwabu, G. and Sandefur, J., 2018. Experimental evidence on scaling up education reforms in Kenya. *Journal of Public Economics*, *168*, pp.1-20.

Brumley, J., 2017. *The 15 All-Time Best-Selling Prescription Drugs, Kiplinger, 5 December*. Available at: www.kiplinger.com/slideshow/investing/t027-s001-the-15-all-time-best-selling-prescription-drugs/index.html (Accessed: 2 July 2023)

Bundesgesundheitsministerium, 2023a. *Mitglieder und Versicherte der gesetzlichen Krankenversicherung (GKV)*. Available at: www.bundesgesundheitsministerium.de/themen/krankenversicherung/zahlen-und-fakten-zur-krankenversicherung/mitglieder-und-versicherte.html (Accessed 22 July 2023)

Bundesgesundheitsministerium, 2023b. *Zuzahlung*. Available at: www.bundesgesundheitsministerium.de/zuzahlung-krankenversicherung.html (Accessed 23 July 2023)

buzer.de, 2023. *Änderung §84 SGB V vom 01.05.2006.* Available at: https://www.buzer.de/gesetz/2497/al820-0.htm (Accessed: 27 June 2023)

Castanheira, M., Ornaghi, C. and Siotis, G., 2019. The unexpected consequences of generic entry. *Journal of Health Economics*, *68*, p.102243.

CDC, 2023. *Coronary Artery Disease (CAD).* Available at: www.cdc.gov/heartdisease/coronary_ad.htm (Accessed 19 July 2023).

Chapman, S.R., Fitzpatrick, R.W. and Aladul, M.I., 2017. Has cost inhibited the uptake of more potent statins in England?. *Pharmacoepidemiology and Drug Safety*, *26*(8), pp.984-991.

Chen, Y. and Li, H., 2009. Mother's education and child health: Is there a nurturing effect?. *Journal of health economics*, *28*(2), pp.413-426.

Ciancio, A., Kämpfen, F., Kohler, H.P. and Kohler, I.V., 2021. Health screening for emerging non-communicable disease burdens among the global poor: Evidence from sub-Saharan Africa. *Journal of health economics*, *75*, p.102388.

Clarke, P.M. and Fitzgerald, E.M., 2010. Expiry of patent protection on statins: effects on pharmaceutical expenditure in Australia. *Medical Journal of Australia*, *192*(11), pp.633-636.

Cohen, J. and Dupas, P., 2010. Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment. *The Quarterly Journal of Economics*, *125*(1), pp.1-45.

Cohen, J., Dupas, P. and Schaner, S., 2015. Price subsidies, diagnostic tests, and targeting of malaria treatment: evidence from a randomized controlled trial. *American Economic Review*, *105*(2), pp.609-645.

Cook, T.D. and Campbell, D.T., 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Settings. Boston: Houghton Mifflin.*

Coscelli, A., 2000. The importance of doctors' and patients' preferences in the prescription decision. *The Journal of Industrial Economics*, *48*(3), pp.349-369.

Cunha, F., Heckman, J.J., Lochner, L. and Masterov, D.V., 2006. Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, *1*, pp.697-812.

Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidharan, K. and Sundararaman, V., 2013. School inputs, household substitution, and test scores. *American Economic Journal: Applied Economics*, *5*(2), pp.29-57.

De Walque, D., 2007. Does education affect smoking behaviors?: Evidence using the Vietnam draft as an instrument for college education. *Journal of health economics*, *26*(5), pp.877-895.

Dennis, J., Ramsay, T., Turgeon, A.F. and Zarychanski, R., 2013. Helmet legislation and admissions to hospital for cycling related head injuries in Canadian provinces and territories: interrupted time series analysis. *BMJ*, *346*.

Department of Health & Social Care, 2023. *Consultation outcome: Proposed update to the 2023 statutory scheme to control the costs of branded health service medicines.* Available at:

www.gov.uk/government/consultations/proposed-update-to-the-2023-statutory-scheme-to-control-the-costs-of-branded-health-service-medicines/proposed-update-to-the-2023-statutory-scheme-to-control-the-costs-of-branded-health-service-medicines (Accessed 3 August 2023).

Derde, L.P., Cooper, B.S., Goossens, H., Malhotra-Kumar, S., Willems, R.J., Gniadkowski, M., Hryniewicz, W., Empel, J., et al., 2014. Interventions to reduce colonisation and transmission of antimicrobial-resistant bacteria in intensive care units: an interrupted time series study and cluster randomised trial. *The Lancet infectious diseases*, *14*(1), pp.31-39.

Duflo, E., 2001. Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American economic review*, *91*(4), pp.795-813.

Duflo, E., Hanna, R. and Ryan, S.P., 2012. Incentives work: Getting teachers to come to school. *American economic review*, *102*(4), pp.1241-1278.

Duflos, G. and Lichtenberg, F.R., 2012. Does competition stimulate drug utilization? The impact of changes in market structure on US drug prices, marketing and utilization. *International Review of Law and Economics*, *32*(1), pp.95-109.

Dupas, P., 2011. Health behavior in developing countries. *Annual Review of Economics*, *3*(1), pp.425-449.

Dylst, P., Vulto, A., and Simoens, S., 2011, The impact of reference-pricing systems in Europe: a literature review and case studies. *Expert Review of Pharmacoeconomics & Outcomes Research,* 11(6), pp. 729-737.

Egan, A. and Colman, E., 2011. Weighing the Benefits of High-Dose Simvastatin against the Risk of Myopathy. *The New England Journal of Medicine, 365*(4), pp. 285-287.

EU, 2023. *Patents*. Available at: europa.eu/youreurope/business/running-business/intellectual-property/patents/index_en.htm (Accessed 24 July 2023)

Firpo, S. and Possebom, V., 2018. Synthetic control method: Inference, sensitivity analysis and confidence sets. *Journal of Causal Inference*, *6*(2).

Fiorentini, G., Bruni, M.L. and Mammi, I., 2022. The same old medicine but cheaper: The impact of patent expiry on physicians' prescribing behaviour. *Journal of Economic Behavior & Organization*, *204*, pp.37-68.

Fischer, K.E. and Stargardt, T., 2016. The diffusion of generics after patent expiry in Germany. *The European Journal of Health Economics*, *17*, pp.1027-1040.

Fleming, K.A., Horton, S., Wilson, M.L., Atun, R., DeStigter, K., Flanigan, J., Sayed, S., Adam, P., et al., 2021. The Lancet Commission on diagnostics: transforming access to diagnostics. *The Lancet*, *398*(10315), pp. 1997-2050.

Flood, D., Seiglie, J.A., Dunn, M., Tschida, S., Theilmann, M., Marcus, M.E., Brian, G., Norov, B., et al., 2021. The state of diabetes treatment coverage in 55 low-income and middle-income countries: a cross-sectional study of nationally representative, individual-level data in 680 102 adults. *The Lancet Healthy Longevity*, *2*(6), pp.e340-e351.

Garcia-Moreno, V., Gertler, P.J. and Patrinos, H.A., 2019. School-based management and learning outcomes: Experimental evidence from Colima, Mexico. *World Bank Policy Research Working Paper*, (8874).

Geldsetzer, P., Manne-Goehler, J., Marcus, M.E., Ebert, C., Zhumadilov, Z., Wesseh, C.S., Tsabedze, L., Supiyev, A., et al., 2019. The state of hypertension care in 44 low-income and middle-income countries: a cross-sectional study of nationally representative individual-level data from 1· 1 million adults. *The Lancet*, *394*(10199), pp.652-662.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., Chang, S.M. and Grantham-McGregor, S., 2014. Labor market returns to an early childhood stimulation intervention in Jamaica. *Science*, *344*(6187), pp.998-1001.

Gertler, P., Patrinos, H. and Rubio-Codina, M., 2012. Empowering parents to improve education: Evidence from rural Mexico. *Journal Of Development Economics*, *99*(1), pp. 69-79.

GKV-Spitzenverband, 2023a. *Die gesetzlichen Krankenkassen*. Available at: www.gkv-spitzenverband.de/krankenversicherung/kv_grundprinzipien/alle_gesetzlichen_krankenkassen/alle_gesetzlichen_krankenkassen.jsp (Accessed 4 July 2023).

GKV-Spitzenverband, 2023b. *Arzneimittel-Festbeträge*. Available at: www.gkv-spitzenverband.de/krankenversicherung/arzneimittel/arzneimittel_festbetraege/festbetraege.jsp (Accessed 20 June 2023).

Glewwe, P., Kremer, M. and Moulin, S., 2009. Many children left behind? Textbooks and test scores in Kenya. *American Economic Journal: Applied Economics*, *1*(1), pp.112-135.

Glewwe, P. and Maïga, E. W., 2011. The impacts of school management reforms in Madagascar: do the impacts vary by teacher type?. *Journal of Development Effectiveness*, *3*(4), pp. 435-469.

Glewwe, P. and Muralidharan, K., 2016. Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. In *Handbook of the Economics of Education* (Vol. 5, pp. 653-743). Elsevier.

Global Education Evidence Advisory Panel, 2020. *Cost-Effective Approaches to Improve Global Learning. What does recent evidence tell us are "Smart Buys" for improving learning in low- and middle-income countries?*

Grabowski, H., Long, G. and Mortimer, R., 2014. Recent trends in brand-name and generic drug competition. *Journal of medical economics*, *17*(3), pp.207-214.

Greene, W.H., 2012. *Econometric Analysis – Seventh Edition*, Pearson Education Limited, Edinburgh.

Güneş, P.M., 2016. The impact of female education on teenage fertility: Evidence from Turkey. *The BE journal of economic analysis & policy*, *16*(1), pp.259-288.

Hamory, J., Miguel, E., Walker, M., Kremer, M. and Baird, S., 2021. Twenty-year economic impacts of deworming. *Proceedings of the National Academy of Sciences*, *118*(14), p.e2023185118.

Haqqi, T., 2023. 25 Most Prescribed Medication in the World, *Yahoo Finance*, 18 May. Available at: finance.yahoo.com/news/25-most-prescribed-medication-world-112510893.html?soc_src=social-sh&soc_trk=ma (Accessed 2 July 2023)

Herkert, D., Vijayakumar, P., Luo, J., Schwartz, J.I., Rabin, T.L., DeFilippo, E. and Lipska, K.J., 2019. Cost-related insulin underuse among patients with diabetes. *JAMA internal medicine*, *179*(1), pp.112-114.

Hooley, B., Afriyie, D.O., Fink, G. and Tediosi, F., 2022. Health insurance coverage in low-income and middle-income countries: progress made to date and related changes in private and public health expenditure. *BMJ Global Health*, *7*(5), p.e008722.

Houses of Parliament, 2010. *Postnote Drug Pricing*. Available at: www.parliament.uk/globalassets/documents/post/postpn_364_Drug_Pricing.pdf (Accessed 3 August 2023).

IFA, 2023. *IFA-Redaktionskalender*. Available at: www.ifaffm.de/de/ifa-fuer-anbieter/ifa-redaktionskalender.html (Accessed 23 July 2023)

IHME. 2023. *GBD Results.* Available at: vizhub.healthdata.org/gbd-results/ (Accessed 17 July 2023).

Imai, S., Fushimi, K. and Sundell, K.A., 2018. Impact of new efficacy information on sales of antihypertensive medicines in Japan and Sweden. *Health Policy and Technology*, *7*(2), pp.194-199.

International Diabetes Federation, 2023. *IDF Diabetes Atlas*. Available at: https://diabetesatlas.org (Accessed 3 August 2023).

Jones, P.J. and Schoeller, D.A., 1990. Evidence for diurnal periodicity in human cholesterol synthesis. *Journal of lipid research*, *31*(4), pp.667-673.

Kaul, A., Klößner, S., Pfeifer, G. and Schieler, M., 2015. Synthetic control methods: Never use all pre-intervention outcomes together with covariates. MPRA Paper No. 83790

Kassenärztliche Bundesvereinigung, 2007. Rahmenvorgaben nach §84 Absatz 7 SGB V und Vereinbarung nach §84 Abs. 7a SGB V – Arzneimittel – für das Jahr 2007. *Deutsches Ärzteblatt,* 104(1-2), pp. A69-72.

Kassenärztliche Bundesvereinigung, 2011. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2012. *Deutsches Ärzteblatt,* 108(47), pp. A2565-A2569.

Kassenärztliche Bundesvereinigung, 2021. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2022. *Deutsches Ärzteblatt,* 118(42), pp. A1961-A1968.

Khattri, N., Ling, C. and Jha, S., 2012. The effects of school-based management in the Philippines: an initial assessment using administrative data. *Journal Of Development Effectiveness*, *4*(2), pp. 277-295.

Klose, G. and Schwabe, U., 1999. Lipidsenkende Mittel. *Arzneiverordnungs-Report 1998*, pp.356-365.

Klose, G. and Schwabe, U., 2001. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2000*, pp.444-454.

Klose, G. and Schwabe, U., 2002. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2001*, pp.483-493.

Klose, G. and Schwabe, U., 2003. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2002*, pp.517-529.

Klose, G. and Schwabe, U., 2004a. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2003*, pp.570-583.

Klose, G. and Schwabe, U., 2004b. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2004*, pp.641-658.

Klose, G. and Schwabe, U., 2006. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2005*, pp.698-715.

Klose, G. and Schwabe, U., 2007. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2006*, pp.690-708.

Klose, G. and Schwabe, U., 2008a. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2007*, pp.651-666.

Klose, G. and Schwabe, U., 2008b. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2008*, pp.645-660.

Klose, G. and Schwabe, U., 2009. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2009*, pp.637-652.

Klose, G. and Schwabe, U., 2010. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2010*, pp.665-680.

Klose, G. and Schwabe, U., 2011. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2011*, pp.683-698.

Klose, G. and Schwabe, U., 2012. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2012*, pp.697-712.

Klose, G. and Schwabe, U., 2013. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2013*, pp.687-702.

Klose, G. and Schwabe, U., 2014. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2014*, pp.733-746.

Klose, G. and Schwabe, U., 2015. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2015*, pp.745-756.

Klose, G. and Schwabe, U., 2016. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2016*, pp.523-530.

Klose, G. and Schwabe, U., 2017. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2017*, pp.529-538.

Klose, G. and Schwabe, U., 2018. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2018*, pp.577-586.

Klose, G. and Schwabe, U., 2019. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2019*, pp.749-560.

Klose, G. and Schwabe, U., 2020. Lipidsenkende Mittel. *Arzneiverordnungs-Report 2020*, pp.601-614.

King, E.M. and Behrman, J.R., 2009. Timing and duration of exposure in evaluations of social programs. *The World Bank Research Observer*, *24*(1), pp.55-82.

Kreif, N., Grieve, R., Hangartner, D., Turner, A.J., Nikolova, S. and Sutton, M., 2016. Examination of the synthetic control method for evaluating health policies with multiple treated units. *Health economics*, *25*(12), pp.1514-1528.

Lakdawalla, D. and Philipson, T., 2012. Does intellectual property restrict output? An analysis of pharmaceutical markets. *The Journal of Law and Economics*, *55*(1), pp.151-187.

Larreguy, H. and Marshall, J., 2017. The effect of education on civic and political engagement in nonconsolidated democracies: Evidence from Nigeria. *Review of Economics and Statistics*, *99*(3), pp.387-401.

Lassibille, G., Tan, J.P., Jesse, C. and Van Nguyen, T., 2010. Managing for results in primary education in Madagascar: Evaluating the impact of selected workflow interventions. *The World Bank Economic Review*, *24*(2), pp.303-329.

Lau, W.C., Murray, M., El-Turki, A., Saxena, S., Ladhani, S., Long, P., Sharland, M., Wong, I.C. and Hsia, Y., 2015. Impact of pneumococcal conjugate vaccines on childhood otitis media in the United Kingdom. *Vaccine*, *33*(39), pp.5072-5079.

Law, M. and Rudnicka, A.R., 2006. Statin safety: a systematic review. *The American Journal of Cardiology,* 97: pp. 52C-60C.

Linden, A. and Adams, J.L., 2011. Applying a propensity score-based weighting model to interrupted time series data: improving causal inference in programme evaluation. *Journal of Evaluation in Clinical Practice*, 17: pp. 1231-1238.

Mach, F., Baigent, C., Catapano, A.L., Koskinas, K.C., Casuala, M., Badimon, L., Chapman, M.J. De Backer, G.G., et al., 2020. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. *European Heart Journal*, 41, pp. 111-88.

Mankiw, N.G., Romer, D. and Weil, D.N., 1992. A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, *107*(2), pp.407-437.

Manne-Goehler, J., Geldsetzer, P., Agoudavi, K., Andall-Brereton, G., Aryal, K.K., Bicaba, B.W., Bovet, P., Brian, G., et al., 2019. Health system performance for people with diabetes in 28 low-and middle-income countries: a cross-sectional study of nationally representative surveys. *PLoS medicine*, *16*(3), p.e1002751.

Manne-Goehler, J., Theilmann, M., Flood, D., Marcus, M.E., Andall-Brereton, G., Agoudavi, K., Arboleda, W.A.L., Aryal, K.K., et al., 2022. Data Resource Profile: The Global Health and Population Project on Access to Care for Cardiometabolic Diseases (HPACC). *International Journal of Epidemiology*, *51*(6), pp. e337-e349.

Marcus, M.E., Ebert, C., Geldsetzer, P., Theilmann, M., Bicaba, B.W., Andall-Brereton, G., Bovet, P., Farzadfar, F., et al., 2021. Unmet need for hypercholesterolemia care in 35 low-and middle-income countries: A cross-sectional study of nationally representative surveys. *PLoS Medicine*, *18*(10), p.e1003841.

Marcus, M.E., Manne-Goehler, J., Theilmann, M., Farzadfar, F., Moghaddam, S.S., Keykhaei, M., Hajebi, A., Tschida, S., et al., 2022. Use of statins for the prevention of cardiovascular disease in 41 low-income and middle-income countries: a cross-sectional study of nationally representative, individual-level data. *The Lancet Global Health*, *10*(3), pp.e369-e379.

Marwick, C., 2003. Bayer is forced to release documents over withdrawal of cerivastatin. *BMJ*, 326(7388): 518.

McGuire, T. and Bauhoff, S., 2011. Adoption of a cost-saving innovation: Germany, UK and Simvastatin. In: Klusen, N., Verheyen, F. and Wagner, C. ed. *England and Germany in Europe–What lessons can we learn from each other*. Baden-Baden, Germany: Nomos Verlag, pp.11-26.

Ministry of Health and Social Affairs, 1982. *The Health and Medical Services Act (1982:763).* Available in English at: www.ilo.org/dyn/travail/docs/1643/health%20a%20nd%20medical%20insurance%20act.pdf (Accessed 3 August 2023).

Mitze, T., Kosfeld, R., Rode, J. and Wälde, K., 2020. Face masks considerably reduce COVID-19 cases in Germany. *Proceedings of the National Academy of Sciences*, *117*(51), pp.32293-32301.

Montenegro, C.E. and Patrinos, H., 2014. Comparable estimates of returns to schooling around the world. *World Bank Policy Research Working Paper*, *7020*, pp.1-41.

Moon, J.C. and Bogle, R.G., 2006. Switching statins. *Bmj*, *332*(7554), pp.1344-1345.

Muralidharan, K. and Sundararaman, V., 2011. Teacher performance pay: Experimental evidence from India. *Journal of political Economy*, *119*(1), pp.39-77.

Nambiar, D., Bhaumik, S., Pal, A. and Ved, R., 2020. Assessing cardiovascular disease risk factor screening inequalities in India using Lot Quality Assurance Sampling. *BMC Health Services Research*, *20*, pp.1-13.

National Audit Office, 2018. *Investigation into NHS spending on generic medicines in primary care.* Available at: www.nao.org.uk/wp-content/uploads/2018/06/Investigation-into-NHS-spending-on-generic-medicines-in-primary-care.pdf (Accessed 3 August 2023).

National Board of Health and Welfare, 2008. *Nationella riktlinjer för hjärtsjukvård 2008*
*Beslutsstöd för prioriteringar.* Available at: https://docplayer.se/1764461-Nationella-riktlinjer-for-hjartsjukvard-2008-beslutsstod-for-prioriteringar.html (Accessed 3 August 2023)

National Population Commission, 2016. *2015 Nigeria Education Data Survey (NEDS)*. Washington DC: United States Agency for International Development.

NCD-RisC, 2023*. Data Downloads*. Available at: ncdrisc.org/data-downloads.html (accessed 3 August 2023).

NCD Risk Factor Collaboration, 2021. Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. *The Lancet*, *398*(10304), pp. 957-980.

Newey, W.K. and West, K.D., 1987. A Simple, Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica,* 55(3), pp. 703-708.

Newman, J., Pradhan, M., Rawlings, L. B., Ridder, G., Coa, R., and Evia, J. L., 2002. An impact evaluation of education, health, and water supply investments by the Bolivian Social Investment Fund. *The World Bank Economic Review*, *16*(2), pp. 241-274.

NHS, 2023a. *Cardiovascular disease*. Available at: www.nhs.uk/conditions/cardiovascular-disease/ (Accessed 19 July 2023).

NHS, 2023b. *NHS prescription charges.* Available at: www.nhs.uk/nhs-services/prescriptions-and-pharmacies/nhs-prescription-charges/ (Accessed 3 August 2023).

NHSBSA, 2021. *Prescription Cost Analysis (PCA) data.* Available at www.nhsbsa.nhs.uk/prescription-data/dispensing-data/prescription-cost-analysis-pca-data (Accessed 21 July 2023).

NICE, 2008. *Cardiovascular disease: identifying and supporting people most at risk of dying early*. London: National Institute for Health and Clinical Excellence. Available at: www.nice.org.uk/guidance/ph15/chapter/1-Recommendations (Accessed 3 August 2023).

NICE, 2014. Cardiovascular disease: risk assessment and reduction, including lipid modification. London: National Institute for Health and Care Excellence. Available at: www.nice.org.uk/guidance/cg181 (Accessed 3 July 2023).

Nigerian National Bureau of Statistics, 2020. *Nigeria Living Standards Survey 2018/2019*. Abuja. Available at: nigerianstat.gov.ng/elibrary?queries[search]=living%20standards%20survey (Accessed 27 October 2020).

Nigerian National Bureau of Statistics and UNICEF, 2017. *Multiple Indicator Cluster Survey 2016-17 Survey Findings Report*. Abuja: National Bureau of Statistics and United Nations Children's Fund. Available at: https://www.unicef.org/nigeria/reports/multiple-indicator-cluster-survey-2016-17-mics (Retrieved 27 October 2020).

Ochmann, S., Owolabi, K.E., Olatunji-David, F., Okunlola, N. and Vollmer, S., 2022. The impact of grants in combination with school-based management trainings on primary education: a cluster-randomized trial in Northern Nigeria. *Journal of Development Effectiveness*, *14*(3), pp.189-208.

Ochmann, S., von Polenz, I., Marcus, M.E., Theilmann, M., Flood, D., Agoudavi, K., Aryal, K.K., Bahendeka, S. et al., 2023. Diagnostic testing for hypertension, diabetes, and hypercholesterolaemia in low-income and middle-income countries: a cross-sectional study of data for 994 185 individuals from 57 nationally representative surveys. *The Lancet Global Health, 11(9), pp.e1363-e1371.*

Office for Health Improvement and Disparities, 2023. *NHS entitlements: migrant health guide.* Available at: www.gov.uk/guidance/nhs-entitlements-migrant-health-guide (Accessed 3 August 2023).

Office for National Statistics, 2022. Estimates of the population for the UK, England, Wales, Scotland and Northern Ireland. Available at: www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland (Accessed 4 July 2023).

Office of Fair Trading, 2007. *Annexe A: Markets for prescription pharmaceuticals in the NHS.* Available at: webarchive.nationalarchives.gov.uk/ukgwa/20140402162831/http://oft.gov.uk/OFTwork/markets-work/pprs (Accessed 3 August 2023).

Olken, B. A., Onishi, J., and Wong, S., 2014. Should aid reward performance? Evidence from a field experiment on health and education in Indonesia. *American Economic Journal: Applied Economics*, *6*(4), pp. 1-34.

Opatrny, M., 2021. The impact of the Brexit vote on UK financial markets: a synthetic control method approach. *Empirica*, *48*(2), pp.559-587.

Our World in Data, 2023. *Measuring progress towards the Sustainable Development Goals.* Available at: sdg-tracker.org (Accessed 20 July 2023).

Peiris, D., Ghosh, A., Manne-Goehler, J., Jaacks, L.M., Theilmann, M., Marcus, M.E., Zhumadilov, Z., Tsabedze, L., et al., 2021. Cardiovascular disease risk profile and management practices in 45 low-income and middle-income countries: A cross-sectional study of nationally representative individual-level survey data. *PLoS medicine*, *18*(3), p.e1003485.

Pettersson, B., Hoffmann, M., Wändell, P. and Levin, L.Å., 2012. Utilization and costs of lipid modifying therapies following health technology assessment for the new reimbursement scheme in Sweden. *Health policy*, *104*(1), pp.84-91.

Pharmazeutische Zeitung, 2012. Atorvastatin: Erstmals als Generikum verfügbar. Available at: www.pharmazeutische-zeitung.de/2012-03/atorvastatin-erstmals-als-generikum-verfuegbar/# (Accessed 27 June 2023)

Piper, B., Destefano, J., Kinyanjui, E. M. and Ong'ele, S., 2018. Scaling up successfully: Lessons from Kenya's Tusome national literacy program. *Journal of Educational Change*, *19*(3), pp. 293-321.

Pradhan, M. P., and De Ree, J., 2014. District Governance and Student Learning in Indonesia. *Asian Development Bank Economics Working Paper Series*, (397).

Pritchett, L. and Beatty, A., 2012. The Negative Consequences of Overambitious Curricula in Developing Countries. *Center for Global Development Working Paper,* (293).

Reibling, N., Ariaans, M. and Wendt, C., 2019. Worlds of healthcare: a healthcare system typology of OECD countries. *Health Policy*, *123*(7), pp.611-620.

Reinikka, R. and Svensson, J., 2011. The power of information in public services: Evidence from education in Uganda. *Journal of Public Economics*, *95*(7-8), pp.956-966.

Roser, M., 2020. *Why is life expectancy in the US lower than in other rich countries?* Our World In Data. Available at: ourworldindata.org/us-life-expectancy-low (Accessed 3 August 2023).

Ruan, Y., Guo, Y., Zheng, Y., Huang, Z., Sun, S., Kowal, P., Shi, Y. and Wu, F., 2018. Cardiovascular disease (CVD) and associated risk factors among older adults in six low-and middle-income countries: results from SAGE Wave 1. *BMC public health*, *18*(1), pp.1-13.

Santibañez, L., Abreu-Lastra, R. and O'Donoghue, J.L., 2014. School based management effects: Resources or governance change? Evidence from Mexico. *Economics of Education Review*, *39*, pp.97-109.

Savin, N.E. and White, K.J., 1977. The Durbin-Watson test for serial correlation with extreme sample sizes or many regressors. *Econometrica: Journal of the Econometric Society*, pp.1989-1996.

Scandinavian Simvastatin Survival Study Group, 1994. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *The Lancet*, *344*(8934), pp.1383-1389.

Schachter, M., 2005. Chemical, pharmacokinetic and pharmacodynamic properties of statins: an update. *Fundamental & clinical pharmacology*, *19*(1), pp.117-125.

SEARCH (Study of the Effectiveness of Additional Reductions in Cholesterol and Homocysteine) Collaborative Group, 2010. Intensive lowering of LDL cholesterol with 80 mg versus 20 mg simvastatin daily in 12 064 survivors of myocardial infarction: a double-blind randomised trial. *Lancet, 376,* pp. 1658-1669.

Selvaraj, S., Farooqui, H.H. and Mehta, A., 2019. Does price regulation affect atorvastatin sales in India? An impact assessment through interrupted time series analysis. *BMJ open*, *9*(1), p.e024200.

Sen, A.K., 1997. Human capital and human capability. *World development*, *25*(12), pp.1959-1961.

Simvastatin over the counter, 2005. *Drug and Therapeutics Bulletin,* 43, pp. 25-28.

Smith, A., 1776. *The Wealth of Nations: An Inquiry into the Nature and Causes of the Wealth of Nations.*

Socialstyrelsen, 2023. *Statistikdatabas för läkemedel*. Available at: sdb.socialstyrelsen.se/if_lak/val.aspx. (Accessed 21 July 2023).

Stargardt, T., 2010. The impact of reference pricing on switching behaviour and healthcare utilisation: the case of statins in Germany. *The European Journal of Health Economics*, *11*, pp.267-277.

Stone, N.J., Robinson, J.G., Lichtenstein, A.H., Bairey Merz, N., Blum, C.B., Eckel, R.H., Goldberg, A.C., Gordon, D., et al., 2013. 2013 ACC/AHA Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults. *Circulation*, 129(suppl 2): pp. S1-S45.

Teufel, F., Seiglie, J.A., Geldsetzer, P., Theilmann, M., Marcus, M.E., Ebert, C., Arboleda, W.A.L., Agoudavi, K., et al., 2021. Body-mass index and diabetes risk in 57 low-income and middle-income countries: a cross-sectional study of nationally representative, individual-level data in 685 616 adults. *The Lancet*, *398*(10296), pp. 238-248.

TLV, 2009. Genomgången av läkemedel vid blodfettrubbningar. Available at: www.tlv.se/download/18.467926b615d084471ac33ed5/1510316365043/slutrapport-blodfett.pdf (Accessed 3 August 2023).

TLV, 2017. *A Brief introduction to the Swedish System for Pricing and Reimbursement of Pharmaceutical Products.* Available at: www.ourcommons.ca/Content/Committee/421/HESA/Brief/BR8900114/br-external/DentalAndPharmaceuticalBenefitsAgency-e.pdf (Accessed 3 August 2023).

TLV, 2023. *PPRI Pharma Profile – Sweden 2023.* Available at: www.tlv.se/download/18.3bcae8a518631688689dc1/1676623001299/ppri_pharma_profile_sweden_2023.pdf (Accessed 3 August 2023).

Toth, F., 2020. Integration vs separation in the provision of health care: 24 OECD countries compared. *Health Economics, Policy and Law*, *15*(2), pp.160-172.

UNESCO, 2019. *UNESCO Institute of Statistics*. Available at: data.uis.unesco.org (Accessed 27 October 2020)

United Nations, 2022. *Department of Economic and Social Affairs Population Division – World Population Prospects 2022.* Available at: population.un.org/wpp/Download/Standard/MostUsed/ (Accessed 4 July 2023).

United Nations, 2023. *The 17 Goals.* Available at: sdgs.un.org/goals (Accessed 20 July 2023).

University of Southern California Dornsife Center for Economic and Social Research, 2023. *Gateway to Global Aging Data*. Available at: g2aging.org (accessed 3 August 2023).

USAID, 2023. *The DHS Program: Demographic and Health Surveys*. Available at: dhsprogram.com (accessed 3 August 2023).

US Food and Drug Administration, 2011. FDA Drug Safety Communication: New restrictions, contraindications, and dose limitations for Zocor (simvastatin) to reduce the risk of muscle injury. Available at: https://www.fda.gov/drugs/drug-safety-and-availability/fda-drug-safety-communication-new-restrictions-contraindications-and-dose-limitations-zocor (Accessed 3 July 2023).

Vondeling, G.T., Cao, Q., Postma, M.J. and Rozenbaum, M.H., 2018. The impact of patent expiry on drug prices: a systematic literature review. *Applied health economics and health policy*, *16*, pp.653-660.

Wagner, A.K., Soumerai, S.B., Zhang, F. and Ross-Degnan, D., 2002. Segmented regression analysis of interrupted time series studies in medication use research. *Journal of clinical pharmacy and therapeutics*, *27*(4), pp.299-309.

Weng, T.C., Yang, Y.H.K., Lin, S.J. and Tai, S.H., 2010. A systematic review and meta-analysis on the therapeutic equivalence of statins. *Journal of clinical pharmacy and therapeutics*, *35*(2), pp.139-151.

WHO, 2018a. *HEARTS Technical package for cardiovascular disease management in primary health care: risk based CVD management*. Available at: apps.who.int/iris/bitstream/handle/10665/260421/WHO-NMH-NVI-18.2-eng.pdf (accessed 3 August 2023).

WHO, 2018b. Pricing of cancer medicines and its impacts. *World Health Organization: Geneva, Switzerland*.

WHO, 2020. WHO Package of Essential Noncommunicable (PEN) Disease Interventions for Primary Health Care. Available at: apps.who.int/iris/rest/bitstreams/1301957/retrieve (accessed 3 August 2023).

WHO, 2022. *Noncommunicable diseases.* Available at: www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases (Accessed 20 July 2023).

WHO, 2023. *WHO Study on global AGEing and adult health (SAGE).* Available at: www.who.int/data/data-collection-tools/study-on-global-ageing-and-adult-health (accessed 3 August 2023).

WHOCC, 2018. *Definition and general considerations*. Available at: https://www.whocc.no/ddd/definition_and_general_considera/ (Accessed: 16 May 2023).

Wilson, M.L., Atun, R., DeStigter, K., Flanigan, J., Fleming, K.A., Horton, S., Kleinert, S. and Sayed, S., 2019. The Lancet Commission on diagnostics: advancing equitable access to diagnostics. *The Lancet*, *393*(10185), pp. 2018-2020.

World Bank, 1993. World Development Report 1993: *Investing in health*. Washington DC: World Bank Group.

World Bank, 2015. *Project Appraisal Document on a Proposed Global Partnership for Education Grant of US$100 Million to the Federal Republic of Nigeria for a Nigeria Partnership for Education Project*. Available at: documents1.worldbank.org/curated/en/506841476077511270/pdf/PAD634-PAD-P143842-Box396300B-PUBLIC-ACS.pdf (Accessed 27 October 2020).

World Bank, 2017. *World Development Report 2018: Learning to realize education's promise*. Washington DC: World Bank Group.

World Bank, 2020. *Human Capital Project*. Available at: www.worldbank.org/en/publication/ human-capital (Accessed 21 September 2020).

World Bank, 2021. *Implementation Completion and Results Report TF-18918 on a Global Partnership for Education Grant in the amount of SDR 69.36 Million (US$100 Million Equivalent) to the Federal Republic of Nigeria for the Nigeria Partnership for Education Project (NIPEP).* Available at: documents1.worldbank.org/curated/en/740801614005370353/pdf/Nigeria-Partnership-for-Education-Project.pdf (Accessed 27 October 2020).

World Bank, 2023a. *DataBank - World Development Indicators*. Available at: databank.worldbank.org/reports.aspx?source=world-development-indicators (Accessed 20 July 2023).

World Bank, 2023b. *World Bank Country and Lending groups.* Available at datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups (Accessed 21 July 2023).

Xu, K., Evans, D.B., Kawabata, K., Zeramdini, R., Klavus, J. and Murray, C.J., 2003. Household catastrophic health expenditure: a multicountry analysis. *The Lancet*, *362*(9378), pp.111-117.

Yadav, H., Shah, D., Sayed, S., Horton, S. and Schroeder, L.F., 2021. Availability of essential diagnostics in ten low-income and middle-income countries: results from national health facility surveys. *The Lancet Global Health*, *9*(11), pp.e1553-e1560.

Yamauchi, F. 2014. *An alternative estimate of school-based management impacts on students' achievements: evidence from the Philippines*. The World Bank.

Zhang, F., Wagner, A.K. and Ross-Degnan, D., 2011. Simulation-based power calculation for designing interrupted time series analyses of health policy interventions. *Journal of clinical epidemiology*, *64*(11), pp.1252-1261.

# 6. Appendix

# 6.1.Appendix for Essay 1

*Appendix table A1.1: ANCOVA and Differences-in-Differences estimates for outcomes of interest*

| | ANCOVA<br>(1) | Diff-in-Diff<br>(2) |
|---|---|---|
| **Infrastructure and equipment** | | |
| Quality of the learning environment[1] | 0.047<br>(0.028) | 0.063<br>(0.041) |
| Sanitation[1] | 0.084*<br>(0.048) | 0.126*<br>(0.068) |
| School facilities[1] | -0.006<br>(0.030) | 0.001<br>(0.048) |
| Any toilets | 0.232***/‡<br>(0.073) | 0.275**<br>(0.111) |
| **Pupil enrolment** | | |
| Total enrolment | 23.96<br>(29.87) | 6.93<br>(61.0) |
| Female enrolment rate | -0.012<br>(0.044) | 0.016<br>(0.723) |
| **Pupil attendance** | | |
| Any students present | -0.048<br>(0.084) | 0.051<br>(0.106) |
| **Teacher attendance** | | |
| Any teacher present | -0.048<br>(0.091) | 0.109<br>(0.123) |
| Any learning taking place | 0.020<br>(0.081) | 0.176<br>(0.110) |
| **Normalised learning achievement scores**[2] | | |
| Grade 2 numeracy | -0.043<br>(0.094) | |
| Grade 2 school means numeracy | | 0.015<br>(0.141) |
| Grade 3 numeracy | -0.028<br>(0.111) | |
| Grade 3 school means numeracy | | 0.098<br>(0.157) |
| Grade 2 school means literacy | -0.016<br>(0.075) | -0.015<br>(0.146) |
| Grade 3 literacy | -0.116<br>(0.108) | |
| Grade 3 school means literacy | | -0.021<br>(0.159) |

[1]These variables are indices ranging from 0 to 1 with 0 representing a poor and 1 a good outcome.

[2]We were unable to match students from baseline to endline, so in column (1) the baseline control variables for the learning achievement scores are baseline school-level averages and in column (2) Differences-in-Differences were estimated using school level averages at both base- and endline.

Note: Standard errors in parentheses. Statistical significance levels (10%, 5%, 1%) based on naïve p-values represented with */**/*** and based on Benjamini-Hochberg adjusted p-values corresponding to a 5% significance level represented with ‡. Sample sizes: School-level regressions (infrastructure and equipment, pupil enrolment, both attendance measures, and all school means literacy) had between 118 and 128 observations. Student-level regressions (normalized learning achievement scores) had between 2021 and 2770 observations. Exception: Female enrolment rate had 56 observations. Also note columns (1) and (3) of Table 2.

*Appendix table A1.2: Baseline balance of normal and high grant treatment groups*

| | Normal | | High | | Difference in means | p-value of difference in means |
|---|---|---|---|---|---|---|
| | N (1) | Mean (2) | N (3) | Mean (4) | (N-H) (5) | (6) |
| **Summary indices** | | | | | | |
| Quality of the learning environment[1] | 31 | 0.338 | 33 | 0.309 | 0.030 | 0.43 |
| Sanitation[1] | 31 | 0.151 | 33 | 0.222 | -0.072 | 0.29 |
| School facilities[1] | 31 | 0.144 | 33 | 0.154 | -0.010 | 0.84 |
| Any toilets | 31 | 0.161 | 33 | 0.152 | 0.010 | 0.92 |
| **Pupil enrolment** | | | | | | |
| Total enrolment | 31 | 177 | 33 | 248 | 71 | 0.34 |
| Female enrolment rate | 25 | 0.392 | 29 | 0.390 | 0.002 | 0.95 |
| **Pupil attendance** | | | | | | |
| No. students observed in Grade 2 / 3 | 31 | 50 | 33 | 51 | -0.126 | 0.99 |
| Pupil absence rate according to registry | 25 | 0.528 | 29 | 0.295 | 0.233 | 0.02** |
| **Teacher attendance** | | | | | | |
| Number of teachers present | 31 | 0.129 | 33 | 0.970 | -0.841 | 0.02** |
| Any learning taking place | 31 | 0.097 | 33 | 0.212 | -0.115 | 0.21 |
| **Learning achievement scores (out of 20)** | | | | | | |
| Grade 2 math | 829 | 6.09 | 872 | 6.22 | -0.126 | 0.48 |
| Grade 3 math | 541 | 8.56 | 622 | 8.97 | -0.411 | 0.13 |
| Grade 2 literacy | 829 | 1.14 | 872 | 1.85 | -0.711 | 0.00***/‡ |
| Grade 2 school means literacy | 30 | 1.11 | 32 | 1.78 | -0.666 | 0.14 |
| Grade 3 literacy | 541 | 3.22 | 622 | 4.11 | -0.891 | 0.00***/‡ |
| Grade 3 school means literacy | 31 | 2.97 | 32 | 3.58 | -0.608 | 0.33 |

[1]These variables are indices ranging from 0 to 1 with 0 representing a poor and 1 a good outcome.
Statistical significance levels (10%, 5%, 1%) based on naïve p-values represented with */**/*** and based on Benjamini-Hochberg adjusted p-values corresponding to a 5% significance level represented with ‡.

*Appendix table A1.3: ITT estimates of normal and high grant treatment groups*

| | Normal SIG (1) | High SIG (2) | p-value of F-Test (1) = (2) (3) |
|---|---|---|---|
| **Summary indices** | | | |
| Quality of the learning environment[1] | 0.033 | 0.038 | 0.90 |
| | (0.039) | (0.038) | |
| Sanitation[1] | 0.038 | 0.099 | 0.37 |
| | (0.060) | (0.059) | |
| School facilities[1] | -0.020 | 0.002 | 0.66 |
| | (0.043) | (0.042) | |
| Any toilets[2] | 0.048 | 0.227** | 0.13 |
| | (0.101) | (0.100) | |
| **Pupil enrolment** | | | |
| Total enrolment | 12.82 | 89.87* | 0.19 |
| | (51.06) | (50.52) | |
| Female enrolment rate | 0.041 | -0.019 | 0.29 |
| | (0.051) | (0.049) | |
| **Pupil attendance** | | | |
| Any students present | 0.097 | -0.210** | 0.009*** |
| | (0.101) | (0.100) | |
| **Teacher attendance** | | | |
| Number of teachers present | -0.4 | 0.429 | 0.97 |
| | (0.736) | (0.828) | |
| Any learning taking place[2] | 0.065 | -0.039 | 0.36 |
| | (0.098) | (0.097) | |
| **Normalised learning achievement scores** | | | |
| Grade 2 numeracy | -0.192* | 0.042 | 0.09* |
| | (0.116) | (0.123) | |
| Grade 3 numeracy | -0.082 | 0.020 | 0.52 |
| | (0.144) | (0.154) | |
| Grade 2 school means literacy | -0.094 | -0.052 | 0.78 |
| | (0.135) | (0.127) | |
| Grade 3 school means literacy | -0.118 | -0.178 | 0.72 |
| | (0.144) | (0.144) | |

[1]These variables are indices ranging from 0 to 1 with 0 representing a poor and 1 a good outcome.

[2]These variables were not balanced at baseline; the reported ITT OLS result is from running an ANCOVA specification.

Note: Initially, balance in outcomes between normal and high treatment schools was tested and confirmed (see Table A2). Standard errors in parentheses. Statistical significance levels (10%, 5%, 1%) based on naïve p-values represented with */**/*** and based on Benjamini-Hochberg adjusted p-values corresponding to a 5% significance level represented with ‡.

Sample sizes: School-level regressions (infrastructure and equipment, pupil enrolment, both attendance measures, and literacy) had between 121 and 125 observations. Student-level regressions (numeracy scores) had between 2021 and 2770 observations. Exception: Female enrolment rate had 65 observations and number of teachers present had 33 observations.

## 6.2.Appendix for Essay 2

### Table of Contents for Appendix 6.2

## Appendix 2.1: Inclusion process for STEPS surveys

Inclusion criteria for a survey:

1) The survey was conducted during or after 2010; in cases where two surveys were available for a particular country, the most recent survey was used;
2) The survey data were made available at the individual level;
3) The survey was conducted in an upper-middle, lower-middle or low-income country according to the World Bank at the time the survey was conducted;
4) The survey was nationally representative;
5) The survey had a response rate ≥30%;
6) The survey contained one or more questions of whether a respondent had been screened for hypertension, diabetes, or hypercholesterolemia.

Data collection process:
"We first identified all countries in which a World Health Organization (WHO) Stepwise Approach to Surveillance (STEPS) survey had been conducted during a year in which the country fell into an eligible World Bank country income category of low-income or middle-income. Prior to the STEPS surveys being made available in the WHO STEPS survey Central Data Catalog in 2019, we systematically requested each eligible STEPS survey from a list of these surveys that the WHO maintains on their website. The research team contacted the responsible party for each survey, based on the information provided on this website. If the contact information was outdated or unavailable, the authors relied on publications utilizing STEPS data and electronic searches of the survey or contact name. For the Caribbean region, country involvement was facilitated by the Caribbean Public Health Agency (CARPHA).

In 2019, additional eligible surveys were downloaded from the Central Data Catalog. The search words used in the WHO Central Data Catalog were: (1) STEPS collection, (2) surveys conducted ≥2008, (3) low-and middle-income countries."[1]

STEPS surveys included: 2016 Algeria, 2017 Azerbaijan, 2018 Bangladesh, 2016 Belarus, 2015 Benin, 2014 Bhutan, 2014 Botswana, 2013 Burkina Faso, 2010 Cambodia, 2011 Comoros, 2010 Costa Rica, 2018 Ecuador, 2010 Eritrea, 2014 Eswatini, 2010 Gambia, 2016 Georgia, 2016 Guyana, 2016 Iran, 2015 Iraq, 2015 Kenya, 2015 Kiribati, 2013 Kyrgyzstan, 2013 Laos, 2017 Lebanon, 2012 Lesotho, 2011 Liberia, 2013 Moldova, 2013 Mongolia, 2017 Morocco, 2014 Myanmar, 2013 Nepal, 2012 Rwanda, 2013 Samoa, 2015 Solomon Islands, 2014 Sri Lanka, 2013 St. Vincent & the Grenadines, 2015 Sudan, 2016 Tajikistan, 2012 Tanzania, 2014 Timor Leste, 2010 Togo, 2014 Tokelau, 2017 Tonga, 2015 Tuvalu, 2014 Uganda, 2011 Vanuatu, 2015 Vietnam, 2017 Zambia, and 2011 Zanzibar.

## Appendix 2.2: Inclusion process for non-STEPS surveys

Search engine: Google

Search terms: "[country name]" AND ("population-based" OR household) AND ("blood glucose" OR "plasma glucose" OR "blood sugar" OR hemoglobin OR haemoglobin OR A1c OR HbA1c OR A1C OR Hb1c OR Hba1c OR HGBA1C OR "blood pressure" OR hypertension OR hypertensive OR cholesterol OR LDL OR HDL OR lipoprotein OR triglycerides OR triglyceride OR lipid OR lipids)

Number of hits reviewed: Hits reviewed until eligible survey identified, or, in the case of no eligible survey identified, first 50 hits (10 hits per page/5 pages reviewed)

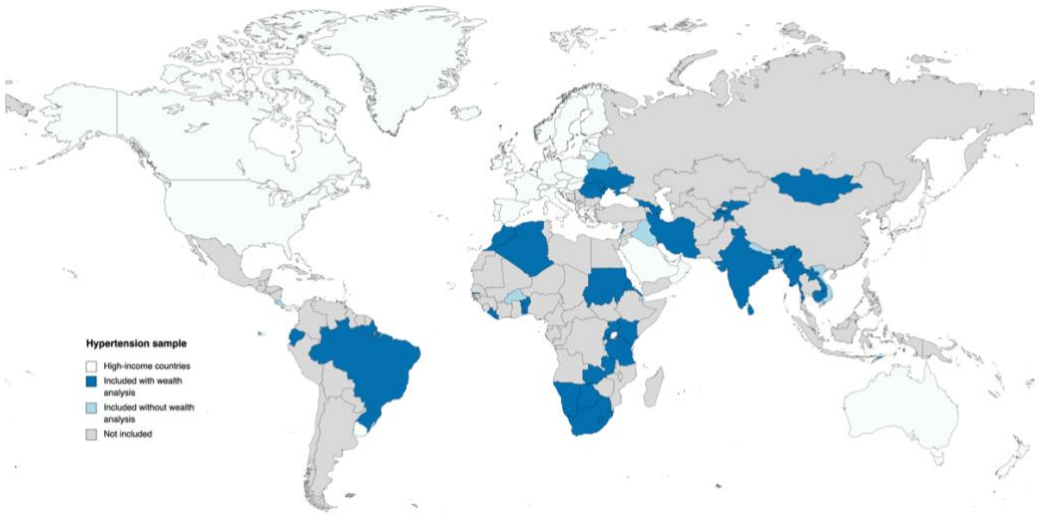
Inclusion criteria for a survey was the same as for STEPS surveys described in Appendix 2.

Countries included in search: Afghanistan, Albania, American Samoa, Angola, Argentina, Armenia, Bangladesh, Belize, Bolivia, Bosnia and Herzegovina, Brazil, Bulgaria, Burundi, Cabo Verde, Cameroon, Central African Republic, Chad, Chile, China, Colombia, Côte d'Ivoire, Cuba, Democratic People's Republic of Korea, Democratic Republic of the Congo, Djibouti, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Ethiopia, Fiji, Gabon, Gambia, Ghana, Grenada, Guatemala, Guinea, Guinea-Bissau, Haiti, Honduras, India, Indonesia, Jamaica, Jordan, Kazakhstan, Kosovo, Madagascar, Malawi, Malaysia, Maldives, Mali, Mauritania, Mauritius, Mexico, Micronesia (Federated States of), Montenegro, Mozambique, Namibia, Nauru, Nicaragua, Niger, Nigeria, North Macedonia, Pakistan, Papua New Guinea, Paraguay, Peru, Philippines, Romania, Russia, Senegal, Serbia, Seychelles, Sierra Leone, Somalia, South Africa, South Sudan, Sri Lanka, St. Lucia, Suriname, Syrian Arab Republic, Thailand, Tunisia, Turkey, Turkmenistan, Ukraine, Venezuela, Yemen, Zimbabwe.

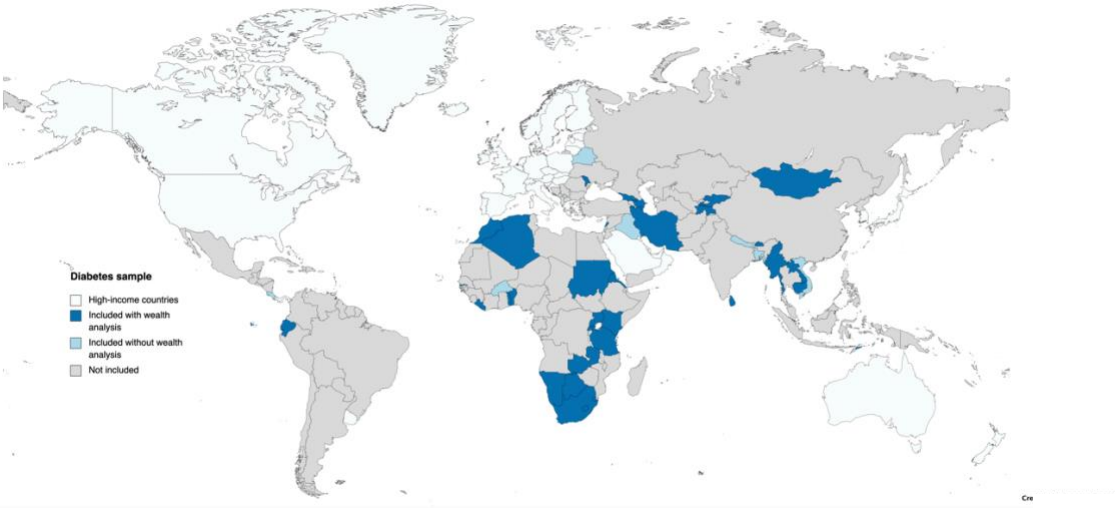
Non-STEPS surveys included:
The 2013 Brazilian Pesquisa Nacional de Saude(PNC), the 2015-2016 Indian National Family Health Survey (NFHS), the 2014 Indonesian Family Life Survey (IFLS), the 2017 Marshall Islands HYBRID Survey, the 2013 Namibia DHS, the 2015-2016 Study for the Evaluation of Prevalence of Hypertension and Cardiovascular Risk in Romania III (SEPHAR), the 2013 Seychelles National Survey of Noncommunicable Diseases, and the 2013 South African National Health and Nutrition Examination Survey (SANHANES).

# Appendix 2.3: Map of included countries

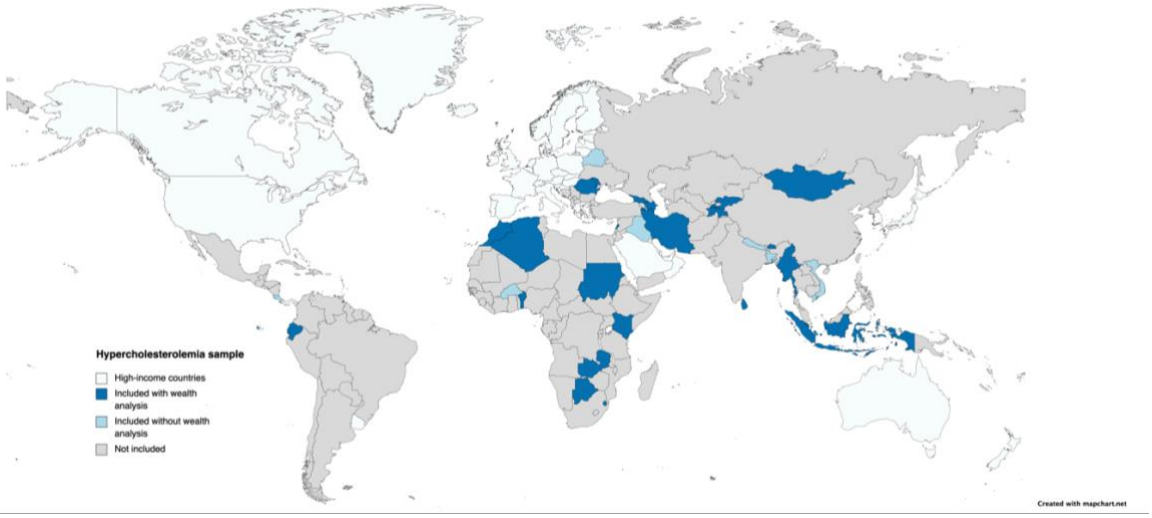
*Appendix figure A2.1: Hypertension analysis sample*




*Appendix figure A2.2: Diabetes analysis sample*




*Appendix figure A2.3: Hypercholesterolaemia analysis sample*



# Appendix 2.4: Survey characteristics

*Appendix table A2.4: Survey characteristics by country*

| Country (1) | World Bank income group (2) | Years of data collection (3) | Sample size (4) | Mean age (5) | Female (in %) (6) | GDP per capita (in int-$) (7) | Missing self-reported diagnostic testing information in … population (in %) * | | | Included in inequality analysis on wealth? (11) |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Hyper-tension (8) | Dia-betes (9) | Hyper-cholesterolemia (10) | |
| Algeria | UMIC | 2016 | 6,956 | 38·2 | 48·4% | 11,511 | 0·0% | 0·0% | 0·0% | Yes |
| Azerbaijan | UMIC | 2017 | 2,801 | 39·6 | 50·4% | 14,402 | 0·0% | 0·0% | 0·0% | Yes |
| Bangladesh | LMIC | 2018 | 8,185 | 36·9 | 49·6% | 4,754 | 0·0% | 0·0% | 0·0% | No |
| Belarus | UMIC | 2016 | 5,010 | 43·0 | 52·2% | 19,249 | 0·0% | 0·0% | 0·0% | No |
| Benin | LIC | 2015 | 5,116 | 34·6 | 47·6% | 3,287 | 0·0% | 0·0% | 0·0% | Yes |
| Bhutan | LMIC | 2014 | 2,812 | 37·7 | 43·0% | 11,832 | 0·0% | 0·0% | 0·0% | Yes |
| Botswana | UMIC | 2014 | 4,055 | 34·2 | 48·0% | 17,767 | 0·0% | 0·0% | 0·0% | Yes |
| Brazil | UMIC | 2013 | 60,202 | 42·9 | 52·4% | 14,652 | 0·0% | 100·0% | 0·0% | Yes |
| Burkina Faso | LIC | 2013 | 4,698 | 39·2 | 52·8% | 2,178 | 13·8% | 0·0% | 0·0% | No |
| Cambodia | LIC | 2010 | 5,433 | 40·4 | 50·6% | 4,389 | 0·0% | 0·0% | 100·0% | Yes |
| Comoros | LIC | 2011 | 5,475 | 40·7 | 69·7% | 3,060 | 0·0% | 0·0% | 100·0% | Yes |
| Costa Rica | UMIC | 2010 | 3,681 | 42·9 | 49·5% | 20,208 | 0·2% | 0·6% | 0·7% | No |
| Ecuador | UMIC | 2018 | 4,638 | 40·1 | 50·5% | 11,375 | 0·0% | 0·0% | 0·0% | Yes |
| Eritrea | LIC | 2010 | 6,265 | 43·4 | 81·0% | NA | 0·0% | 0·0% | 100·0% | Yes |
| Eswatini | LMIC | 2014 | 3,274 | 33·9 | 54·4% | 8,622 | 0·0% | 0·0% | 0·0% | Yes |
| Gambia | LIC | 2010 | 4,090 | 37·7 | 49·5% | 2,223 | 0·0% | 0·0% | 100·0% | Yes |
| Georgia | LMIC | 2016 | 4,204 | 42·8 | 51·5% | 14,993 | 0·0% | 0·0% | 0·0% | Yes |
| Guyana | UMIC | 2016 | 2,655 | 37·5 | 48·0% | 13,082 | 0·0% | 0·0% | 0·0% | Yes |
| India | LMIC | 2015-2016 | 742,842 | 33·0 | 47·0% | 6,700 | 0·0% | 100·0% | 100·0% | Yes |
| Indonesia | LMIC | 2014 | 30,978 | 41·5 | 52·0% | 11,812 | 100·0% | 100·0% | 0·0% | Yes |
| Iran | UMIC | 2016 | 30,032 | 44·4 | 51·5% | 12,389 | 0·1% | 0·1% | 0·2% | Yes |
| Iraq | UMIC | 2015 | 4,071 | 36·5 | 46·4% | 10,881 | 17·0% | 36·9% | 0·0% | No |
| Kenya | LMIC | 2015 | 4,484 | 34·9 | 49·5% | 4,330 | 0·0% | 0·0% | 0·1% | Yes |
| Kiribati | LMIC | 2015 | 2,122 | 38·6 | 53·9% | 2,272 | 0·0% | 0·0% | 0·0% | Yes |
| Kyrgyzstan | LMIC | 2013 | 2,623 | 40·8 | 48·1% | 5,254 | 0·0% | 0·0% | 0·0% | Yes |
| Laos | LMIC | 2013 | 2,541 | 38·4 | 57·1% | 7,826 | 0·0% | 0·0% | 100·0% | Yes |
| Lebanon | UMIC | 2017 | 1,899 | 41·1 | 51·5% | 14,552 | 0·0% | 0·0% | 0·0% | Yes |
| Lesotho | LMIC | 2012 | 2,307 | 38·0 | 49·4% | 2,704 | 0·0% | 0·0% | 100·0% | Yes |
| Liberia | LIC | 2011 | 2,525 | 36·9 | 55·8% | 1,428 | 0·0% | 0·0% | 100·0% | Yes |
| Marshall Islands | UMIC | 2017 | 3,015 | 39·3 | 50·9% | 3,889** | 0·0% | 0·0% | 0·1% | Yes |
| Moldova | LMIC | 2013 | 4,755 | 39·5 | 47·3% | 13,022 | 0·0% | 0·1% | 0·0% | Yes |
| Mongolia | LMIC | 2013 | 6,013 | 40·5 | 50·1% | 12,317 | 0·0% | 0·0% | 0·0% | Yes |
| Morocco | LMIC | 2017 | 5,429 | 41·7 | 49·9% | 7,515 | 0·0% | 0·0% | 0·0% | Yes |
| Myanmar | LMIC | 2014 | 8,266 | 41·8 | 49·2% | 5,142 | 0·0% | 0·0% | 0·0% | Yes |
| Namibia | UMIC | 2013 | 3,679 | 46·7 | 58·8% | 9,637 | 0·3% | 0·7% | 100·0% | Yes |
| Nepal | LMIC | 2019 | 5,593 | 36·7 | 52·9% | 3,417 | 0·0% | 0·0% | 0·0% | No |
| Romania | UMIC | 2015-2016 | 1,970 | 48·0 | 52·6% | 29,873 | 0·0% | 100·0% | 0·0% | Yes |
| Rwanda | LIC | 2012 | 7,223 | 33·4 | 51·5% | 2,227 | 0·0% | 0·0% | 100·0% | Yes |
| Samoa | LMIC | 2013 | 1,765 | 36·8 | 46·4% | 6,521 | 0·0% | 0·0% | 100·0% | Yes |
| Seychelles | UMIC | 2013 | 1,240 | 42·6 | 49·9% | 29,223 | 100·0% | 0·0% | 0·0% | Yes |
| Solomon Islands | LMIC | 2015 | 2,506 | 36·8 | 52·5% | 2,663 | 0·0% | 0·0% | 0·0% | Yes |
| South Africa | UMIC | 2012 | 15,473 | 38·8 | 55·0% | 12,482 | 0·4% | 0·7% | 100·0% | Yes |
| Sri Lanka | LMIC | 2014 | 5,166 | 39·2 | 48·5% | 13,078 | 0·0% | 0·0% | 0·0% | Yes |
| St. Vincent & the Grenadines | UMIC | 2013 | 3,504 | 35·5 | 49·8% | 12,485 | 0·0% | 0·0% | 0·0% | Yes |
| Sudan | LMIC | 2015 | 7,722 | 34·3 | 43·2% | 3,958 | 0·0% | 0·0% | 0·0% | Yes |
| Tajikistan | LMIC | 2016 | 2,718 | 32·0 | 46·5% | 3,380 | 0·0% | 0·0% | 0·0% | Yes |
| Tanzania | LIC | 2012 | 5,545 | 39·0 | 49·4% | 2,660 | 0·0% | 0·0% | 100·0% | Yes |
| Timor Leste | LMIC | 2014 | 2,600 | 41·2 | 57·6% | 3,553 | 0·0% | 0·0% | 0·1% | Yes |
| Togo | LIC | 2010 | 4,311 | 34·2 | 51·8% | 1,597 | 0·0% | 0·1% | 100·0% | Yes |
| Tokelau | UMIC | 2014 | 554 | 35·3 | 52·7% | NA | 0·0% | 0·0% | 0·0% | No*** |
| Tonga | UMIC | 2017 | 3,858 | 40·6 | 63·5% | 6,383 | 0·0% | 0·0% | 0·0% | No |
| Tuvalu | UMIC | 2015 | 1,152 | 37·8 | 51·8% | 4,281 | 0·0% | 0·0% | 0·0% | Yes |
| Uganda | LIC | 2014 | 3,974 | 35·2 | 55·8% | 2,187 | 0·0% | 0·0% | 100·0% | Yes |
| Vanuatu | LMIC | 2011 | 4,639 | 39·6 | 52·1% | 3,153 | 0·0% | 0·5% | 100·0% | Yes |
| Vietnam | LMIC | 2015 | 3,750 | 39·1 | 49·8% | 8,041 | 0·0% | 0·0% | 0·0% | No |
| Zambia | LMIC | 2017 | 4,300 | 33·7 | 49·2% | 3,470 | 0·0% | 0·0% | 0·0% | Yes |
| Zanzibar | LIC | 2011 | 2,491 | 38·8 | 50·7% | 2,836 | 0·0% | 0·0% | 100·0% | Yes |

GDP per capita data are from 2019 shown in constant 2017 international dollars as estimated by the World Bank.

* A missingness of 100% indicates that the country was not included in the respective CVD risk factor sample at all.

** GDP per capita for Marshall Islands is from 2018.

*** Tokelau has missing education data and is excluded from the education analysis.

## Appendix 2.5: Country-specific sampling methods.

*Note: To ensure accuracy in reporting, sampling methods are pasted verbatim from specified sources.*

**Algeria: STEPS 2016-2017**
A multi-stage cluster sample of households. One individual within the age range of the survey was selected per household. Analysis weights were calculated by taking the inverse of the probability of selection of each participant. These weights were adjusted for differences in the age-sex composition of the sample population as compared to the target population.
Different weight variables are available per Step:
wStep1 - for interview data
wStep2 - for physical measures
wStep3 - for biochemical measures

This allows for differences in the weight calculation for each Step of the survey as the age-sex composition of the respondents to each Step can differ slightly due to refusal or drop out. Additionally, some countries perform subsampling for Step 2 and/or Step 3. When no subsampling is done and response rates do not differ across Steps of the survey, the 3 weight variables will be the same.

Age range of participants included: 18-69 years
*Source: no report or fact sheet available. Sampling information obtained from: https://extranet.who.int/ncdsmicrodata/index.php/catalog/91/study-description*

**Azerbaijan: STEPS 2017**
A multi-stage cluster sample of households. One individual within the age range of the survey was selected per household. Analysis weights were calculated by taking the inverse of the probability of selection of each participant. These weights were adjusted for differences in the age-sex composition of the sample population as compared to the target population.

Different weight variables are available per Step:
wStep1 - for interview data
wStep2 - for physical measures
wStep3 - for biochemical measures

This allows for differences in the weight calculation for each Step of the survey as the age-sex composition of the respondents to each Step can differ slightly due to refusal or drop out. Additionally, some countries perform subsampling for Step 2 and/or Step 3. When no subsampling is done and response rates do not differ across Steps of the survey, the 3 weight variables will be the same.

Age range of participants included: 18-69 years
*Source: no report or fact sheet available. Sampling information obtained from: https://extranet.who.int/ncdsmicrodata/index.php/catalog/127/studydescription#page=overview&tab=stu dy-desc*

**Bangladesh: STEPS 2018**
Sampling Procedure: A multistage complex sampling design was used to produce representative data for that age range in Bangladesh. Response Rate: The overall response rate was 83.8%.

Weighting: Analysis weights were calculated by taking the inverse of the probability of selection of each participant. These weights were adjusted for differences in the age-sex composition of the sample population as compared to the target population. Different weight variables are available per Step:
wStep1 - for interview data
wStep2 - for physical measures
wStep3 - for biochemical measures

This allows for differences in the weight calculation for each Step of the survey as the age-sex composition of the respondents to each Step can differ slightly due to refusal or drop out. Additionally, some countries perform subsampling for Step 2 and/or Step 3. When no subsampling is done and response rates do not differ across Steps of the survey, the 3 weight variables will be the same."Age range of participants included: 25 to 69 years

*Source:https://extranet.who.int/ncdsmicrodata/index.php/catalog/770/study-description#page=overview&tab=study-desc*
*Source: National Institute of Population Research and Training (NIPORT), Mitra and Associates, and ICF International. 2013. Bangladesh Demographic and Health Survey 2011. Dhaka, Bangladesh and Calverton, Maryland, USA: NIPORT, Mitra and Associates, and ICF International.*

**Belarus: STEPS 2016**
The sampling frame is a collection of data and materials from which are selected for the survey. The optimal sampling frame should be complete, accurate and current. Best of all, the above criteria are met by the results of the population census, which became the basis for constructing the sample for the STEPS study. Census population represents a representative territorial sampling frame in the form a hierarchical set of parcels grouped in a certain way. Plots censuses are, on average, about the same size. For each site there is a schematic map that provides a clear, non-overlapping demarcation of geographic districts, as well as information on the population and the number of households.
The largest in size is the census area, which includes several instructor sites. The smallest unit in the hierarchical structure of parcels by censuses - enumeration areas.A positive aspect of using enumeration areas as primary sampling units (PSUs) is that they have a small and approximately the same size (each includes about 100 HHs on average). Consequently this, the PSU is a territory within which it is possible to effectively organize field work. To conduct a population census, the territory of the Republic of Belarus was divided into almost 32 thousand enumeration areas. Due to the fact that the last population census in the Republic of Belarus was carried out in 2009, to update the sample, the current data of polyclinics were used, medical outpatient clinics, FAPs and rural Soviet accounting in rural areas.

Age range of participants included: 18-69 years
*Source: Translated directly from the Belarus STEPS 2016 report. Available at: https://extranet.who.int/ncdsmicrodata/index.php/catalog/100/related_materials*

**Benin: STEPS 2015**

The STEPS survey on risk factors for non-communicable diseases in Benin was conducted from October to December 2015. It was a population-based survey of adults aged 18 to 69 years. A 3-stage sampling frame was used to produce representative data for this age group in Benin. The information required for the investigation was collected electronically using a manual device. The survey was implemented by the National Program for the Fight against Non-Communicable Diseases (PNLMNT) of the Ministry of Health of Benin. A total of 5,126 adults participated in the STEPS survey conducted in Benin. The overall response rate was 98.6%. The 1st survey took place in 2008. A third survey is planned for 2020 if the financial situation allows it.

Age range of participants included: 18-69 years

*Source: Translated directly from the Benin STEPS 2015 report. Available at: https://extranet.who.int/ncdsmicrodata/index.php/catalog/107/download/1044*

**Bhutan: STEPS 2014**

To achieve a nationally representative sample, a multistage sampling method was used to select enumeration areas, households and eligible participants at each of the selected households in three stages. The 2005 National Census was chosen as the basis for the sampling frame, with “Geogs” (blocks) in rural areas and towns in urban areas forming the primary sampling units (PSUs). Since the population distribution for urbanicity is 70:30 (rural:urban), 63 PSUs in rural and 14 PSUs in urban areas were chosen. PSUs were selected through the probability proportionate to size (PPS) sampling using the number of households in each PSU. Two secondary sampling units (SSUs) for every rural PSU and 4 SSUs for every urban PSU were selected. This led to the selection of 126 SSUs from rural and 56 SSUs from urban areas. This was also carried out by PPS sampling, using the number of households in each SSU. A total of 16 households from each SSU (both rural and urban) were selected using systematic random sampling. The sampling frame for this was the list of households with a unique identification number (ID) developed by the enumerators for the survey. At the household level, the Kish sampling method was used to randomly select one eligible member (aged 18–69 years) of the household for the survey. The Kish method ranks eligible household members in order of decreasing age, starting with males and then females, and randomly selects a respondent using the automated program for Kish selection in the handheld personal digital assistant (PDA).

*Source: National survey for noncommunicable disease risk factors and mental health using approach WHO Steps Approach in Bhutan – 2014 Available at: http://www.who.int/chp/steps/bhutan/en/.*

**Botswana: STEPS 2014**

Botswana has a population of over 2 million with 27 districts and 4,845 enumeration areas and sample size of 300 enumeration areas with a target population of 6,400 people was

systematically drawn from a pool of the whole enumeration areas. Against the identified enumeration areas numbers of households were listed and proportion of participants was calculated from the total sample size required for the country. Finally a computer generated random number was drawn to go into specific households in that specific enumeration area and at the end eligible participants residing in the household were listed into the electronic hand held data assistant (PDA) and at the end a name was picked automatically to participate in the survey.
Age range of participants included: 15-69 years
*Source: Botswana STEPS report. Available at:*
*https://extranet.who.int/ncdsmicrodata/index.php/catalog/318*

**Brazil: Pesquisa Nacional de Saúde 2013**
The text below was translated from: https://www.pns.icict.fiocruz.br/index.php?pag=planoamostragem

"The Master Sample is a set of units of areas that are selected to meet various surveys of the IBGE Integrated System of Household Searches (SIPD). These units are considered primary sampling units (PSUs) in the sample planning of each of the surveys that use the Master Sample, such as PNS. The sampling plan consists of the stratification of the UPAs and selection of these units with probability proportional to the size, given by the number of permanent private households (DPPs).

The register for selection of the Master Sample was a file containing information from the Demographic Census 2010 on the census tracts of the geographic scope, whose limits are defined in the Operational Geographic Base 2010, totaling 316574 sectors. A sector or set of sectors with at least 60 DPPs was defined as UPA, with the exception of a few units, because it was not possible to aggregate sectors in some municipalities.

The stratification of the UPAs obeys four different criteria: administrative, including the division of the UF into capital, rest of the Metropolitan Region (RM) or Integrated Region of Economic Development - RIDE, and rest of the UF; geographical subdivision, which subdivides capitals and other large municipalities into more strata; situation that involves rural / urban categorization; and the statistician in order to improve the accuracy of the estimates.

As part of the SIPD, the sampling design of the PNS followed, in part, the sampling design of the Master Sample, especially with regard to the stratification of the UPAs.

The PNS sample is by clusters in three stages of selection:
- 1st stage: selection with probability proportional to the size (given by the number of DPPs in each unit) of the UPAs sub-sample in each stratum of the Master Sample;
- 2nd stage: selection by simple random sample of households in each UPA selected in the first stage;
- 3rd stage: selection by simple random sampling of the adult (person aged 18 years or older) among all adult residents of the household.

The PNS will integrate the SIPD, which will make it possible to relate the information collected with other researches, such as the PNAD and the Household Budget Survey (POF) at different levels of geographic aggregation."

*Pesquisa Nacional de Saúde. Plano de Amostragem. 2010. https://www.pns.icict.fiocruz.br/index.php?pag=planoamostragem. Accessed May 11, 2018.*

**Burkina Faso: STEPS 2013**

Sampling methodology: The study was conducted on a sample obtained from a three-stage cluster stratified as recommended by the WHO for STEPS screening surveys. risk factors for noncommunicable diseases.

The sampling frame used was that derived from the general census of the population and habitat 2006 (RGPH 2006) and updated in 2010 during the survey Demographic and Health Survey of Burkina Faso (EDS-BF, 2010). This update concerned the enumeration areas (EAs) that correspond to the cluster as part of this study.

Selection of clusters: The choice of clusters was made according to a systematic random selection proportional to their size (in number of households) within strata (regions). To do this clusters were organized by stratum and place of residence (urban / rural). A total of 240 clusters of which 185 were in rural areas and 55 in urban areas were selected for the investigation.

Selection of households: Households were randomly drawn after an enumeration exhaustive list of all households in the cluster. A draw tool designed on Excel by the team. The technique was used in the field for selecting households to investigate. In total, 20 households in clusters were selected to participate in the study.

Selection of individuals: The choice of individuals was made randomly using Kish's method. In total, an individual aged 25 to 64 living in a selected household was fired for participate in the survey.

*Source, translated from: Rapport de l'enquete national sur la prevalence des principaux facteurs de risques communs aux maladies non transmissibles au Burkina Faso Enquete STEPS 2013. Available at: http://www.who.int/chp/steps/burkina_faso/en/.*

**Cambodia: STEPS 2010**

The initial planned sample size was designed to involve 5,760 persons in accordance with the NCD multi-stage cluster survey method (1.5 design effect, 95% confidence interval, 5% margin or error, and 50% baseline levels of the indicators) in order to provide an equivalent distribution of the participants in regards to age groups and gender after taking into consideration that the estimated potential rate for non- response in each group and refusals in the nest stages would equal to 20%. Estimates were obtained for each of the following eight age/sex groups: men ahed 25-34 years, 35-44 years, 45-54 years, and 55-64 years; and women aged 25-34 years, 35-44 years, 45-54 years, and 55-64 years.

The survey was designed to cover all geographical areas of Cambodia and a 3-stage sampling process as part of the multi-stage cluster sampling was carried out to randomly select the target population: random selection of communes (Khum in rural areas and its equivalent Sangkat in urban area) as primary sampling unit (PSU), followed by villages (Phum) for the

second sampling unit (SSU), and by households for the elementary units (EU). Finally, all members of the randomly chose households aged 25-64 years were invited to participate in this survey. The selection process was performed identically for urban and rural areas in order to get a self-weighted estimate for the whole population of the country. A total of 180 clusters with 34 clusters from the urban area and 146 clusters from the rural area were randomly selected.
Age range of participants included: 25-64 years
*Source: Cambodia STEPS 2010 survey report. Available at: https://www.who.int/ncds/surveillance/steps/cambodia/en/*

**Comoros: STEPS 2011**
The STEPS survey on risk factors for chronic diseases in the Union of the Comoros took place from January to March 2011. This study has undertaken Step 1, Step 2 and Step 3. Indeed, socio-demographic and behavioral measures were collected in Step 1. Physical measures such as height, weight and tension were collected in Step 2 and biochemical measurements were collected to assess the levels of blood glucose and cholesterol levels in Step 3. The STEPS survey conducted in Comoros Union is a survey of general population, targeting adults aged 25 to 64 years. A stratified survey was used to produce representative data for this age group. A total of 5556 adults aged 25 to 64 participated in the STEPS survey on a sample of 5760 people representing an overall response rate of 96.5%.
*Source, translated from* Union des Comores STEPS 2011 Note de synthèse.
*Available at: http://www.who.int/chp/steps/comoros/en/.*

**Costa Rica: STEPS 2010**
The Costa Rican NCRFSS survey was a cross-sectional survey based on a probabilistic cluster sampling design. The NCRFSS survey was conducted during 2010 under the supervision of the Caja Costarricense de Seguro Social, a government public healthcare provider, and covers the overall adult population aged ≥20 years. Multistage cluster sampling was performed stratified by geographical areas, age groups (20–39, 40–64, and ≥65 years) and gender. The first sample stage was the randomized selection of the country's geographical areas as primary sample units followed by the random selection of sectors in selected areas as secondary sample units. The random selection of areas and sectors was performed with probability proportional to size; the area or sector size was determined by the population >20 years during 2009, as estimated by the Costa Rican Census and Statistics National Institute (INEC). Households were chosen through a random number generator using dwelling lists obtained from the health technician assistant in every community until all age group and gender strata sample sizes were achieved. A family dwelling was defined as a group of people who share the same table to eat. Survey participants were selected by the Kish method, which samples participants within a household with equal probability of selection, as recommended by the WHO STEPwise methodology. To be eligible for inclusion in the study, subjects had to be ≥20 years of age, permanently residing in the selected homes, and to have provided written consent. Pregnant or lactating mothers and those who were within 6 months postpartum were excluded from the study. Each participant selected for the study was informed of the study objectives and details before agreeing to participate in the investigation. In all, 3653

noninstitutionalized adults were surveyed, with an 87.8% response rate of the eligible population.
*Source: Wong-McClure R, Gregg EW, Barcelo A, Sanabria-Lopez L, Lee K, Abarca-Gomez L, Cervantes-Loaiza M, Luman ET. Prevalence of diabetes and impaired fasting glucose in Costa Rica: Costa Rican National Cardiovascular Risk Factors Survey, 2010. J Diabetes. 2016 Sep;8(5):686-92.*

**Ecuador STEPS 2018**
Type and stages of the sample design. The STEPS sample was selected following an element probability sampling scheme with the following three stages of selection: i) first stage: selection of Primary Sampling Units (PSU) per stratum; ii) second stage: selection of 12 occupied households within each PSU selected in the first stage; and, iii) third stage: selection of 1 person between 18 and 69 years old per household. Study domains. Men and women between 18 and 69 years of age at the national level, with the exception of Galapagos.
Sample selection. The selection of the PSUs, according to the established size, was carried out independently in a random manner in each of the strata. Twelve households were also randomly selected from each previously selected cluster. From the second survey period onwards, given the high rates of occupancy change, 16 dwellings per conglomerate were selected to counteract this effect. The change affected the remaining 230 clusters, giving a total of 6,680 dwellings to be surveyed. Finally, a list was made of the persons eligible for selection within each dwelling, randomly selecting one of them.
Age range of participants included: 18-69 years
*Source: Ecuador STEPS 2019 Report [Translated]. Available at: https://extranet.who.int/ncdsmicrodata/index.php/catalog/774/study-description#page=sampling&tab=study-desc*

**Eritrea: STEPS 2010**
A multi-stage cluster sample of households. One individual within the age range of the survey was
selected per household.
Analysis weights were calculated by taking the inverse of the probability of selection of each participant. These weights were adjusted for differences in the age-sex composition of the sample population as compared to the target population.
Different weight variables are available per Step:
wStep1 - for interview data
wStep2 - for physical measures
wStep3 - for biochemical measures
This allows for differences in the weight calculation for each Step of the survey as the age-sex composition of the respondents to each Step can differ slightly due to refusal or drop out. Additionally, some countries perform subsampling for Step 2 and/or Step 3. When no subsampling is done and response rates do not differ across Steps of the survey, the 3 weight variables will be the same.
Age range of participants included: 25-74 years

*Source: no report available. Sampling information obtained from: https://extranet.who.int/ncdsmicrodata/index.php/catalog/589/study-description#page=sampling&tab=study-desc*

**Eswatini: STEPS 2014**

A Multi-stage cluster sampling design was applied. The survey covered all the four regions of the country. The size of the country and the distances between the regions and communities made it possible for the survey to sample a population representing all the 4 regions. The Multi-stage sampling procedure was implemented in the following procedural steps:

Stage 1: All four regions were included as a sampling frame of our Primary Sampling Unit (PSU).The number of the PSUs at this stage ensured precision in the survey estimates and as a result 216 PSUs were selected using probability proportional to size sampling.

Stage 2: The second stage of cluster sampling procedure entailed listing, sorting and random systematic sampling of the Secondary Sampling Units (Households) within the PSUs selected in stage1 where 20 households were selected from each PSU. Based on census data, only households with eligible participants were systematically sampled through random systematic sampling.

Stage 3: At this level, all the eligible participants within a household were sequentially listed into the PDAs and only one participant per household was randomly sampled using KISH method built into the PDAs. The KISH method is a widely used technique that uses a pre-assigned table of random numbers to identify the person to be interviewed.
*Source: WHO STEPS: Noncommunicable Disease Risk Factor Surveillance Report Swaziland 2014. Available at: http://www.who.int/chp/steps/swaziland/en/.*

**Gambia: STEPS 2010**

Geographic coverage: national
Age range of participants included: 25-64 years
Step 1 and Step 2 were carried out.
Sampling procedure: A multi-stage cluster sample of households. One individual within the age range of the survey was selected per household.
Overall response rate: 77.9%

Weighting: Analysis weights were calculated by taking the inverse of the probability of selection of each participant. These weights were adjusted for differences in the age-sex composition of the sample population as compared to the target population.
Different weight variables are available per Step:
wStep1 – for interview data
wStep2 – for physical measures
This allows for differences in the weight calculation for each Step of the survey as the age-sex composition of the respondents to each Step can differ slightly due to refusal or drop out. Additionally, some countries perform subsampling for Step 2. When no subsampling is done

and response rates do not differ across Steps of the survey, the 2 weight variables will be the same.
Source: Gambia STEPS 2010, available at: https://extranet.who.int/ncdsmicrodata/index.php/catalog/616#metadata-coverage

**Georgia: STEPS 2016**
The STEPS survey of noncommunicable disease (NCD) risk factors in Georgia was carried out from June 2016 to September 2016. Georgia carried out Step 1, Step 2 and Step 3. Socio demographic and behavioural information was collected in Step 1. Physical measurements such as height, weight and blood pressure were collected in Step 2. Biochemical measurements were collected to assess blood glucose and cholesterol levels in Step 3. The survey was a population-based survey of adults aged 18-69. A Multi-stage cluster sampling design was used to produce representative data for that age range in Georgia. A total of 5554 adults participated in the survey. The overall response rate was 75.7%.
*Source: Georgia STEPS Survey 2016 Fact Sheet.*
*Available at: http://www.who.int/chp/steps/georgia/en/.*

**Guyana: STEPS 2016**
A response rate of 66.68% will be selected based on the experience and response rates of other surveys over the years such as the recent Demographic Health Survey 2009. [...] STEPS 3 involve taking blood samples from a proportion of the sample, in this case 50% of the sample, in order to measure raised blood glucose levels and abnormal blood lipids. [...] The STEPS sample will be prepared by the Bureau of Statistics Guyana following the recommended STEPS sample methodology. A multi-stage cluster sampling design will be used. Guyana is divided into 10 administrative regions and within the administrative regions there are seven towns and each region is further divided into enumeration districts. For the STEPS survey 288 enumeration districts will be selected using the population probability sampling method and from each enumeration district 12 households will be selected giving a total sample size of 3456. Further at the household level each participant will be randomly selected by the electronic tablet. For STEP 3 50% of the sample will be randomly selected to participate. A re-listing of some households may also be necessary, such as those interior region locations, in which case in addition to household listings, enumeration districts maps will also be provided so that a re-listing can be done where required.
*Source: STEPwise Approach to Chronic Disease risk factor surveillance (STEPS): Guyana's Implementation Plan. June 20, 2016. Ministry of Public Health, Guyana.*

**India: NFHS 2015-16**
The NFHS-4 sample was designed to provide estimates of all key indicators at the national and state levels, as well as estimates for most key indicators at the district level (for all 640 districts in India, as of the 2011 Census). The total sample size of approximately 572,000 households for India was based on the size needed to produce reliable indicator estimates for each district and for urban and rural areas in districts in which the urban population accounted for 30-70 percent of the total district population. The rural sample was selected through a two-stage sample design with villages as the Primary Sampling Units (PSUs) at the first stage (selected

with probability proportional to size), followed by a random selection of 22 households in each PSU at the second stage. In urban areas, there was also a two-stage sample design with Census Enumeration Blocks (CEB) selected at the first stage and a random selection of 22 households in each CEB at the second stage. At the second stage in both urban and rural areas, households were selected after conducting a complete mapping and household listing operation in the selected first-stage units.
*Source: Ministry of Health and Family Welfare (MoHFW) - Government of India. India - National Family Health Survey 2015-2016. Report generated on: February 7, 2018.*

**Indonesia: IFLS 2014-15**
Because it is a longitudinal survey, IFLS5 drew its sample from IFLS1, IFLS2, IFLS2+, IFLS3 and IFLS4. The IFLS1 sampling scheme stratified on provinces and urban/rural location, then randomly sampled within these strata (see Frankenberg and Karoly, 1995, for a detailed description). Provinces were selected to maximize representation of the population, capture the cultural and socioeconomic diversity of Indonesia, and be costeffective to survey given the size and terrain of the country. For mainly cost-effectiveness reasons, 14 of the then existing 27 provinces were excluded.3 The resulting sample included 13 of Indonesia's 27 provinces containing 83% of the population: four provinces on Sumatra (North Sumatra, West Sumatra, South Sumatra, and Lampung), all five of the Javanese provinces (DKI Jakarta, West Java, Central Java, DI Yogyakarta, and East Java), and four provinces covering the remaining major island groups (Bali, West Nusa Tenggara, South Kalimantan, and South Sulawesi).
Within each of the 13 provinces, enumeration areas (EAs) were randomly chosen from a nationally representative sample frame used in the 1993 SUSENAS, a socioeconomic survey of about 60,000 households. The IFLS randomly selected 321 enumeration areas in the 13 provinces, over-sampling urban EAs and EAs in smaller provinces to facilitate urban-rural and Javanese–non-Javanese comparisons.

Within a selected EA, households were randomly selected based upon 1993 SUSENAS listings obtained from regional BPS office. A household was defined as a group of people whose members reside in the same dwelling and share food from the same cooking pot (the standard BPS definition). Twenty households were selected from each urban EA, and 30 households were selected from each rural EA. This strategy minimized expensive travel between rural EAs while balancing the costs of correlations among households. For IFLS1 a total of 7,730 households were sampled to obtain a final sample size goal of 7,000 completed households. This strategy was based on BPS experience of about 90% completion rates. In fact, IFLS1 exceeded that target and interviews were conducted with 7,224 households in late 1993 and early 1994.

In IFLS1 it was determined to be too costly to interview all household members, so a sampling scheme was used to randomly select several members within a household to provide detailed individual information.
*Source: Strauss, J., F. Witoelar, and B. Sikoki. "The Fifth Wave of the Indonesia Family Life Survey (IFLS5): Overview and Field Report". March 2016. WR-1143/1-NIA/NICHD.*

**Iran: STEPS 2016**

The sampling part, which includes determining the sample size and the cluster head, belongs to the pre- study phase and was planned in the form of a specific protocol for sample size and statistical sampling. All experts in the quality control team supervised the finding of samples and cluster heads.

In order to estimate the prevalence rate of the risk factors for non-communicable diseases in the country in 1395, a sampling method proportionate to the population was used, which is a common approach in survey studies. Therefore, the selected sample size was proportionated to the population of that province. On the other hand, for estimating the prevalence of the risk factors in the province, in order to be on the safe side, the smallest sample size for achieving the predicted rates was calculated at 95%. This rate was equal to 384 samples, which was selected as the smallest sample size in the least populated province, Ilam. The required sample size for other provinces was therefore calculated according to the population of that province proportionate to the population of the reference province, Ilam. Besides, to control the non- response error, 10% was added to the calculated sample size in each province. In order to decrease costs and increase efficiency, for provinces with 800 samples or more, weights were given to their samples. Weight-giving is an effective method used in surveys in order to decrease the sample size. This was achieved in the selected provinces by considering the calculated sample size as half and the sampling weight as double. The total sample size was calculated to be 30150 and to achieve this sample size, sampling from 3015 clusters was required.

Age range of participants included: 18 and older

*Source: Iran STEPS 2015 report. Available at: https://www.who.int/ncds/surveillance/steps/STEPS_2016_Atlas_EN.pdf?ua=1*

**Iraq: STEPS 2015**

The sample frame consisted of the population of Iraq of (18+) years for both sexes residing in the urban and rural area. It was based on the results of listing and numbering operation for the year 2009 that covered all governorates. Due to the unstable conditions at the time of the survey three governorates (Naynawa, Salahaddin and Al-Anbar) were excluded. A major challenge confronted was the late demographic change due to population movement, displacement and migration. All permanent residents of (18+) years of age, who were resident in Iraq within one month at the time of implementation of the survey were considered eligible.

A cross-sectional community based survey covering 15 governorates in Iraq. A Multi-stage cluster sampling technique was depended to select the minimum representative sample size to estimate the prevalence of the risk factors of noncommunicable disease through direct interview, physical examination and laboratory examination of blood samples of study participants. A total of 412 clusters were randomly selected each contain ten households. One subject from each household was randomly selected using KISH table to participate in the survey with a total sample size of 4120. The Sample was designed to provide estimates on a number of indicators on the situation of Noncommunicable diseases risk factors in Iraq at the national level. A national based rather than a governorate based sample is selected. A multi stage cluster sampling was used with stratification to urban and rural areas. Primary sampling units (PSUs) were the blocks, which consisted of 70 households or more before selection.

Age range of participants included: 18 years and older

*Source: Iraq STEPS 2015 report.*
*Available at: https://www.who.int/ncds/surveillance/steps/Iraq_2015_STEPS_Report.pdf*

**Kenya: STEPS 2015**

The 2015 Kenya STEPs survey was a national cross-sectional household survey designed to provide estimates for indicators on risk factors for non-communicable diseases for persons age 18 – 69 years. The sample was designed with a sample size of 6,000 individuals to allow national estimates by sex (male and female) and residence (urban and rural areas). The survey used the fifth National Sample Surveys and Evaluation Programme (NASSEP V) master sample frame that was developed and maintained by KNBS. The frame was developed using the Enumeration Areas (EAs) generated from the 2009 Kenya Population and Housing Census to form 5,360 clusters split into four equal sub-samples. A three-stage cluster sample design was adopted for the survey involving selection of clusters, households and eligible individuals. In the first stage, 200 clusters (100 urban and 100 rural) were selected from one sub-sample of NASSEP V frame. A uniform sample of 30 households from the listed households in each cluster was selected in the second stage of sampling. The last stage of sampling was done using Personal Digital Assistants (PDAs) at the time of survey, where one individual was randomly selected from all eligible listed household members using a programmed KISH method of sampling.

*Source: WHO: Kenya STEPwise Survey for Non Communicable Diseases Risk Factors 2015 Report. Available at: http://www.who.int/chp/steps/Kenya_2015_STEPS_Report.pdf?ua=1.*

**Kiribati: STEPS 2015**

The second Kiribati STEPS Survey was a population-based survey of 18-69 year olds. The decision was to use three age groups: 18-29, 30-44, 45-69 years for men and women using the following corrections:

- Design Effect of 1.0 (clustering at village and household level)
- 95% confidence interval; p value .05
- 0.7% response rate
- Baseline prevalence percentage indicator: 0.5
- FPC – not applicable
- 6 age-sex groups (18-29 years, 30-44 years, 45-69 years)

As STEPS is intended to be nationally representative, a multi-stage cluster sampling method was used. The STEPS sampling spreadsheet was completed using the most recent census information (2012). The sample was selected in two stages assuming no replacement. At the first stage, a sample of Enumeration Areas (Islands and villages) from each stratum using probability proportional to size (PPS) sampling was selected. In the second stage, a fixed number of households from each selected Enumeration Area using systematic sampling was selected. The third stage of sampling selection was done at the household level using the KISH method.

The sampling identified that data collection would be needed on the following islands: Makin, Butaritari, Marakei, Abaiang, North Tarawa, South Tarawa,Betio, Maiana, Abemama, Kuria, Aranuka, Nonouti, Tabiteuea North, Tabiteuea South, Arorae, Tabuaeran and Kiritimati. Further details in Annex 3.

Age range of participants included: 18 to 69 years

*Source: Kiribati STEPS 2015 report. Available at: https://extranet.who.int/ncdsmicrodata/index.php/catalog/724*

**Kyrgyzstan: STEPS 2013**

A multi-stage cluster sample of households. One individual within the age range of the survey was
selected per household.
Analysis weights were calculated by taking the inverse of the probability of selection of each participant.
These weights were adjusted for differences in the age-sex composition of
the sample population as compared to the target population.
Different weight variables are available per Step:
wStep1 - for interview data
wStep2 - for physical measures
wStep3 - for biochemical measures
This allows for differences in the weight calculation for each Step of the survey as the age-sex composition of the respondents to each Step can differ slightly due to refusal or drop out.
Age range of participants included: 25 to 64 years
*Source: no report or fact sheet available. Sampling information obtained from: https://extranet.who.int/ncdsmicrodata/index.php/catalog/271/study-description#page=overview&tab=study-desc*

**Lao People's Democratic Republic: STEPS 2013**

A multi-stage cluster sample of households. One individual within the age range of the survey was
selected per household. Analysis weights were calculated by taking the inverse of the probability of
selection of each participant. These weights were adjusted for differences in the age-sex composition of the sample population as compared to the target population.
Different weight variables are available per Step:
wStep1 - for interview data
wStep2 - for physical measures
wStep3 - for biochemical measures
This allows for differences in the weight calculation for each Step of the survey as the age-sex composition of the respondents to each Step can differ slightly due to refusal or drop out. Additionally, some countries perform subsampling for Step 2 and/or Step 3. When no subsampling is done and response rates do not differ across Steps of the survey, the 3 weight variables will be the same.
Age range of participants included: 18 to 64 years
*Source: no report or fact sheet available. Sampling information obtained from: https://extranet.who.int/ncdsmicrodata/index.php/catalog/588/study-description#page=sampling&tab=study-desc*

**Lebanon: STEPS 2017**

A national cross-sectional survey adopting a two-stage cluster sampling design was conducted for Steps 1, 2 and 3. The sampling frames references used were the population distribution in Lebanon 2014, retrieved from the Central Administration for Statistics (CAS) and the Syrian population distribution data 2015, retrieved from UNHCR. 144 clusters were selected for the Lebanese sample and 144 clusters for the Syrian sample. The Primary Sampling Units (PSUs) were cadastral areas (cadasters) and the Secondary Sampling Units (SSUs) were the households. Twenty participants were recruited from each cluster. The latest available population estimates (cadastral data) were used, to randomly recruit PSUs by Probability Proportionate to Size (PPS). To account for the issue of the variability in the cadasters' sizes, very small cadasters (<200 individuals) were combined with neighboring PSUs before selecting the sample, to enhance the likelihood of finding 20 target participants. On the other hand, cadasters with a large population size that were guaranteed to be sampled at least twice were handled as strata and each stratum were assigned a fixed number of random starting points based on how often it was selected with certainty. This was done using satellite images divided into grids, previously obtained from the Centers for Disease Control and Prevention (CDC) for all Lebanese cadasters.

For the Lebanese sample, the research team relied on the standard Expanded Program for Immunization (EPI) method for a systematic random selection of the households. Accordingly, within each selected PSU, households were identified using a systematic random approach following the WHO-UNICEF-EPI cluster method. The fieldworkers started with the highest floor on the right side of a building. If the household hosted an eligible participant, they proceeded with data collection, if not, they visited a second household which is selected by skipping 5 households. If during sampling, non-Lebanese households were selected, the fieldworker skipped them in a straight line until a Lebanese household was identified. This method has been previously used for national surveys in Lebanon. One participant was randomly selected within each household, using the eSTEPS application. Households were chosen until the target of 20 participants was reached.

The PSUs for the Syrian refugees' sample were identified, using the most recent available refugee estimates to randomly recruit PSUs by PPS. The same measures aforementioned were done to account for the variation in the cadasters' sizes. The WHO-UNICEF- EPI cluster method was employed to select households. The fieldworkers targeted Syrian households; accordingly, when during sampling, non- Syrian households were selected, the fieldworker skipped them in a straight line until a Syrian household was identified. One participant was randomly selected within each household, using the eSTEPS application.

For both samples, following STEPS' team recommendations, sampling of participants was done without replacement, i.e. once a person was selected that person was not replaced with another one. Efforts were made to include all selected households. If the house was unoccupied at the time of the visit or if an adult was not available for an interview at the time of the visit, that house was revisited up to 4 times, with different visiting times. The number of refusals and non-responses was recorded.

Age range of participants included: 18 to 69 years

*Source: Lebanon STEPS 2016-2017 report. Available at: www.who.int/ncds/surveillance/steps/Lebanon_STEPS_report_2016-2017.pdf?ua=1*

**Lesotho: STEPS 2012**

A multi-stage cluster sample of households. One individual within the age range of the survey was selected per household. Analysis weights were calculated by taking the inverse of the probability of selection of each participant. These weights were adjusted for differences in the age-sex composition of the sample population as compared to the target population.

Different weight variables are available per Step:

wStep1 - for interview data

wStep2 - for physical measures

wStep3 - for biochemical measures

This allows for differences in the weight calculation for each Step of the survey as the age-sex composition of the respondents to each Step can differ slightly due to refusal or drop out. Additionally, some countries perform subsampling for Step 2 and/or Step 3. When no subsampling is done and response rates do not differ across Steps of the survey, the 3 weight variables will be the same.

Age range of participants included: 25-64 years

*Source: Source: no report available. Sampling information obtained from: https://extranet.who.int/ncdsmicrodata/index.php/catalog/491/study-description#page=sampling&tab=study-desc*

**Liberia: STEPS 2011**

Random multi-cluster sampling method was used to collect data during this survey in 5 of the 15 counties of Liberia with the district serving as the primary sampling unit. Different sampling frames were designed and used at the district (Primary Sampling Unit-PSU), Chiefdoms (Secondary Sampling Unit-SSU) and household levels. Households listing generated from the 2008 National Population Census was used, and in each household, the list of individuals' resident was obtained and the Kish Method was used. Kish Method is a household sampling technique developed by WHO for STEPS. The field team selected households by using nutrition sampling method (throwing a pencil to get a selected direction). When the household enumeration sampling point is established, the interviewer counts all the households and using interval sample to get the household number. In each household, one person was selected using the Kish method.

*Source: WHO: The Final Report on the Liberia STEPS Survey 2011. Available at: http://www.who.int/chp/steps/Liberia_2011_STEPS_Report.pdf?ua=1.*

**Marshall Islands: HYBRID 2017**

Stage 1: Households were identified at random according to geographical stratification in Majuro and Ebeye. The country was stratified into two major groups, Urban (Majuro and Ebeye) and Rural (all outer islands). In Majuro and Ebeye, household cluster sampling was used to randomly select households in these areas.

Stage 2: In Majuro and Ebeye, one individual was selected at random from each household using the KISH table method. All adults in Kili, Arno, Wotje, and Jabwor, Jaluit atolls were included in the sample because the adult populations are about 200 each on these atolls.

Age of participants included: ≥18

*Source: Republic of the Marshall Islands Hybrid Survey Final Report 2018. Available at: https://extranet.who.int/ncdsmicrodata/index.php/catalog/742*

**Moldova: STEPS 2013**

A total of 4807 randomly selected respondents participated in the survey. They were all aged 18–69 years, and the group comprised both sexes, as well as residents of all districts and the territorial administrative unit "Gagauz-Yeri", along with Chişinău and Balti municipalities. The survey did not cover the districts from the left bank of the Nistru River and the municipality of Bender. A two-stage cluster sampling procedure was carried out to select randomly participants from among the target population. Cluster sectors from the 2004 Moldova Population Census were used as a basic unit. Given the differences in lifestyle and disease status between populations in urban and rural areas, the target population was stratified into urban and rural areas of residence for the STEPS survey. At the first stage, within each stratum, primary sampling units (PSUs) (enumeration areas (EAs)) were selected systematically with probability proportional to the 2004 Population Census EAs (measure of size equal to the number of population in the EAs, provided by the census). Before selection, the census sectors were sorted geographically from north to south within each stratum, in order to ensure additional implicit stratification according to geographical criteria. A total of 400 clusters representing 400 EAs were selected from the 10 991 census EAs. These probabilistically selected clusters were used also in Moldova's DHS conducted in 2005, and the Multiple Indicator Cluster Surveys (MICS) conducted in 2012. Cartographic materials from the Population Census conducted in Moldova in 2004 were not available, thus it was not possible to use them for the STEPS survey. Therefore, for the first stage the probabilistic samples from the abovementioned surveys were used.

Out of the 400 selected clusters, 167 were rural and 233 were urban. The distribution of the sample of 400 PSUs (EAs) for the DHS/MICS surveys was inversely proportional to the number of population within each stratum, taking into account that the response rate is lower in urban areas than rural owing to the smaller average size of the households in urban areas compared with rural areas. Thus, disproportional allocation with oversampling for urban areas was applied in the STEPS survey. A final weighting adjustment procedure was carried out to enable estimates at national and urban/rural levels.

At the second stage, 15 households (secondary sampling units (SSUs)) were selected within each of the 400 PSUs. From the updated list of households used for the MICS 2012 survey, 15 households were selected randomly per cluster, using the Microsoft Excel® random sample tool. A total of 6000 individuals were selected from among the 400 clusters. The Kish method *(17)* was applied for the random selection of one individual aged 18–69 years from each household.

Age of participants included: 18-69 years

*Source: Republic of Moldova STEPS 2013 report. Available at: https://www.who.int/ncds/surveillance/steps/Moldova_2013_STEPS_Report.pdf*

**Mongolia: STEPS 2013**

A nationwide, cross-sectional survey was conducted covering 8 districts of Ulaanbaatar city and 21 aimags of Mongolia. A total of 6013 individuals aged 15-64 years old, representing the Mongolian adult population, were involved in the survey.

Sampling: The survey was designed to cover all geographical areas of Mongolia, and a multi stage stratified sampling process was carried out to randomly select participants from the target population. Given the urban vs. rural differences in lifestyle and disease status, the target population was stratified into urban and rural areas and the sample was drawn proportionally based on the target population in each area. Ulaanbaatar, Darkhan and Erdenet cities represented urban areas, while the remaining aimags and soums represented rural areas.
Primary units for Ulaanbaatar, Darkhan and Erdenet cities were khoroos, whereas soums served as primary units for rural areas. The same principle used in the previous STEPS surveys in 2005 and 2009 was applied for sampling unit selections for each stage. From each selected household at the tertiary units of multi-stage cluster sampling in both urban and rural areas, only one individual aged 15-64 years old was randomly selected.
The survey covered a total of 65 cluster sampling units. These units included randomly selected individuals from 32 soums in 21 rural aimags and 33 khoroos in Ulaanbaatar, Darkhan and Erdenet cities. The below Table-1 presents selected clusters, cluster sampling units and the numbers and proportion of participants out of the total population. In order to be able to compare the survey results and findings by urban and rural areas, we conducted sampling based on the principles to select approximately similar numbers of participants from both urban and rural areas.
Age of participants included: 15-64 years
*Source: Mongolia STEPS 2013 reports. Available at: extranet.who.int/ncdsmicrodata/index.php/catalog/615/related_materials*

**Morocco: STEPS 2017**
One of the essential elements for establishing a probability sampling plan is the constitution an adequate sampling frame. For the purpose of the STEPS survey, the sampling frame used to meet the sampling need was the 2014 master sample, developed by the HCP based on data from the 2014 population and housing census. It has the advantage extrapolate the sample results to the target population and estimate the accuracy desired. The stratification of observation units belonging to any sampling frame makes it possible to design sampling plans ensuring optimal sample size; a significant reduction in costs and a substantial improvement in the accuracy of expected estimators. However, the choice of criteria allowing the population to be divided into homogeneous groups (strata) and having recent and reliable data on these criteria is a task that requires generally considerable efforts (updating the sampling frame) both in terms of methodological than that of data collection.
In Morocco, the particularity of cities containing several social categories for which, synthesizing the vector of heterogeneous demographic and socioeconomic behavior into a representative characteristic makes stratification a difficult task. The stratification adopted was geographical for the two environments according to the weight in terms of households, each of which has a specific stratification: For urban units, the criteria used were the administrative division into regions, provinces / prefectures and the dominant habitat type. As for the rural environment, the primary units were stratified according to the geographical criterion, and the type of relief dominant at the municipal level.
Age range of participants included: 18 years and older
*Source: Morocco STEPS report [translated online]: https://extranet.who.int/ncdsmicrodata/index.php/catalog/544/study-description*

**Myanmar: STEPS 2014**

To achieve a nationally representative sample, a multi-stage sampling method was used to select townships, wards and villages, households and eligible participants at each of the selected households. *Stage 1: Selection of primary sampling units (PSUs)* Administratively, Myanmar is divided into 330 townships. A township is subdivided into wards for urban settings and village tracts and then villages for rural settings. The list of townships has been used as the sampling frame at the first stage of sampling. Townships form the Primary Sampling Units (PSUs). Out of the total 330 PSUs, 52 PSUs were selected using Probability Proportionate to Size of population in each PSU (PPS). *Stage 2: Selection of Secondary Sampling Units (SSUs)*

From each selected PSU (township), 6 SSUs (wards and villages) were chosen using probability proportionate to population size, totaling 312 SSUs for the whole country. *Stage 3: Selection of eligible participants at household level*

From each selected SSU (ward/village), 30 households were selected using systematic random sampling. The sampling frame for this sampling is the list of households with unique identification number (ID) developed from a recent listing of households available from the Basic Health Staff.

*Stage 4: Selection of eligible participants at household level*

One eligible participant (aged between 25 and 64 years) in the selected households was recruited for the survey. The Kish sampling method was used to randomly select one eligible member of the household. Using the Kish Method, eligible participants (adults aged 25 to 64 years) in each household were ranked in order of 8 decreasing age, starting with males then females, then randomly selected using the automated program for Kish selection in the handheld PDA. Each PSU (township) was estimated to contribute 180 participants, totaling **9,360** participants for 52 selected townships for the whole country. In actual study, the total sample size was 8757 participants.

Age range of participants included: 18 years and older

*Source: STEPwise approach to chronic disease risk factor surveillance report 2014. Available at: https://www.who.int/ncds/surveillance/steps/myanmar/en/*

**Namibia: DHS 2013**

The sample for the 2013 NDHS was a stratified sample selected in two stages. In the first stage, 554 EAs were selected with a stratified probability proportional to size within the sampling frame. The EA size is the number of households residing in the EA and recorded in the 2011 NPHC. Stratification was achieved by separating each region into urban and rural areas. Therefore, the 13 regions were stratified into 26 sampling strata: 13 rural strata, and 13 urban strata. Samples were selected independently in each stratum, with a predetermined number of EAs selected as shown in Table A.3. Implicit stratification with proportional allocation was achieved at each of the lower administrative unit levels by sorting the sampling frame before the sample selection. Sorting was done according to the constituency and the EA code within a sampling stratum, and by using a probability proportional-to-size selection procedure.

After the selection of EAs and before the main survey, a household listing operation was carried out in all selected EAs, and the resulting lists of households served as a sampling frame for the selection of households in the second stage. Some of the selected EAs may large. To limit the amount of work done to list each household, selected EAs with more than 200 households were segmented by the listing team in the field before the household listing. Only one segment was selected for the survey, with probability proportional to the segment size. Household listing was conducted only in the selected segment (see detailed instructions for segmentation in the DHS Manual for Household Listing). So a 2013 NDHS cluster is either an EA or a segment of an EA. In the second-stage selection, a fixed number of 20 households was selected in every urban cluster and rural cluster, by equal probability systematic sampling. A spreadsheet indicating the selected household numbers for each cluster was prepared. The survey interviewers interviewed only the pre-selected households. To prevent bias, no replacements and no changes of the pre-selected households were allowed in the implementing stages. In half of the selected households where there was no male survey, all women age 15-49 were interviewed; in the other half of the selected households where there was a male survey, all males and females age 15-64 were interviewed.
*Source: The Nambia Ministry of Health and Social Services (MoHSS) and ICF International. 2014. The Namibia Demographic and Health Survey 2013. Windhoek, Namibia, and Rockville, Maryland, USA: MoHSS and ICF International.*

**Nepal: STEPS 2013**
The surveyed population included men and women aged 15–69 years who had been living at their place of residence for at least six months. [...]

The sample size was calculated to represent the entre target population in Nepal. In order to achieve this statistical inference, the sample size calculator by WHO (sample_size_calculator STEPS) was used to derive a sample size of 4,200. […]

Probability proportionate to size (PPS) was applied in the sampling strategy to improve the precision of the survey estimates. [...]

For this survey, the Ilaka was taken as the primary sampling unit (PSU). Out of the 921 Ilakas in Nepal, 159 are in the mountains, 467 in the hills and 295 in the Terai. The Steering Committee and the WHO NCD STEPS team at WHO headquarters in Geneva predetermined the number of PSUs to be taken in the study as 70. Thus, 70 Ilakas were sampled. Considering the varied distribution of the population across the ecological belts and to avoid the risk of under selection of the sample from the sparsely populated mountain belt, the distribution of "Ilakas" across ecological belts was determined on the basis of the population distribution pattern in the ecological belts (mountains 7%, hills 43% and Terai 50%). Hence, 30 Ilakas were selected from the hills, 5 from the mountains and 35 from the Terai using PPS. [...]
For the survey, wards (sub-units of VDCs and municipalities) were considered as clusters and taken as the secondary sampling unit (SSU). Three clusters were selected from each of the sampled Ilakas using PPS. All wards for each of the selected Ilakas were listed in order according to their numeric code, then 210 wards were selected (3 wards from each of the 70 Ilakas). To select the three wards from the list, all of the wards in the Ilaka were given a unique

identification number, listed in ascending order along with household size and populated in the software. The software then selected the wards randomly on the basis of PPS.

Twenty households were selected from each cluster using systematic sampling. Thus, a total of 4,200 households were selected from the 210 clusters (20 households per cluster or ward). The sampling interval was determined by dividing the total number of households in the selected wards by 20. [...]
In municipalities, one ward covers a large number of households and each ward has more than 5 and sometimes up to 100 streets (margs or toles). Two margs or toles were selected and ten households were selected from each of the two margs or toles using systematic random sampling. If two or more families were found living in a house, one family was selected randomly. Eligible candidates (15–69 years) from the selected household were listed according to age and sex (males first and then females, in descending order), which was then fed into the Kish program in the personal digital assistants (PDAs), which automatically randomly selected one eligible candidate from each house.
*Source: WHO: Non Communicable Diseases Risk Factors: STEPS Survey Nepal 2013. Available at: http://www.who.int/chp/steps/nepal/en/.*

**Romania: SEPHAR II**
Sampling was performed by a multi-stratified procedure, leading to the selection of a representative sample of 1942 adults. Subject selection followed the principle of equality of chances of being enrolled in the study, regardless of the size of the place of residency.

Stratification criteria for sample selection were:

- territorial regions (Romania's territory was divided into 7 regions plus the capital city Bucharest, based on the National Statistics Institute recommendations: the North-East region, the South-East region, the South region, the South-West region, the West region, the North-West region, the Central region and the Bucharest region);
- locality type (cities with over 200 000 inhabitants, cities with 50 000–200 000 inhabitants, cities with less than 50 000 inhabitants, Commune);
- gender (male and female);
- age groups (18–24 years, 25–34 years, 35–44 years, 45–54 years, 55–64 years, 65–80 years).

In the first stage of selection, the adult population weighted average was calculated for each region and each district, and, based on this, the number of adult persons from each region/district was calculated from the working sample of 1942 subjects.
In the second stage of selection, the number of localities of a certain size from which the subjects were later selected was established for each district. This number was directly proportional to the population in the respective district. A random selection of a certain locality in a certain category was done using a computer software (generation of random numbers). The selected localities represent the interview centers where the study was to take place. The weighted average of the specific locality population in the district was calculated, and, based on this, the number of people selected to participate in the study.
The third stage of selection consisted of distribution by gender of adult people selected from each locality, using Romania's population gender distribution according to the 2002 census (F : M = 51.25% vs. 48.75%) and the fourth stage of selection consisted of distribution by age of

male and female adult people selected from each locality, using Romania's population age distribution according to the 2002 census.
*Source: Dorobantu M, Tautu OF, Darabont R, Ghiorghe S, Badila E, Dana M, Dobreanu M, Baila I, Rutkowski M, Zdrojewski T. Objectives and methodology of Romanian SEPHAR II Survey. Project for comparing the prevalence and control of cardiovascular risk factors in two East-European countries: Romania and Poland. Arch Med Sci. 2015 Aug 12;11(4):715-23.*

**Rwanda: STEPS 2012-2013**
Participants were Rwandan residents aged 15-64 years. Because it was not feasible to conduct a census on the whole population, a representative random sample of participants was selected. To detect statistically significant differences between categories, the WHO STEPwise methodology suggests a minimum sample of 384 people for every age, sex rural/urban or province category the results will be stratified by. For the Rwandan survey the MOH was interested in looking at both males and females across five age groups (15-24 years, 25-34 years, 35-44 years, 45-54 years and 55-64 years), yielding a minimum required sample size of 3840. This was multiplied by 1.5 to account conservatively for the likelihood of a selected participant having the risk factor of interest and then divided by 0.80 assuming that only 80% of those invited to participate would actually participate. This yielded a required sample size of 7200 participants.
Multistage cluster sampling was used to select these participants from the population based on information from the last census. The three levels of clustering were: 1. Random selection of a statistical enumeration area (as defined by NISR) 2. Random selection of a household within the enumeration area 3. Random selection of an individual within the household. Administratively, Rwanda is divided into thirty districts. In turn, each district is subdivided into sectors. Each sector is sub-divided into cells and then into villages. Villages are synonymous with enumeration area's (EAs) in Rwanda and there are a total of 14,953 EAs in Rwanda. A total of 180 EA's (or 1.2%) were randomly selected from this total using a probability proportional to size method that gives those EA's with more people living in them a higher chance of being selected. In this way, the representativeness of the selected EAs is maximized.
Age range of participants included: 15-64 years
*Source: Republic of Rwanda Non-communicable Diseases Risk Factors Report 2012. Available at: https://extranet.who.int/ncdsmicrodata/index.php/catalog/709*

**Samoa: STEPS 2013**
The STEPS survey of chronic disease risk factors in Samoa was carried out from April 2013 to May 2013. Samoa carried out Step 1, Step 2, and Step 3. Socio demographic and behavioural information was collected in Step 1. Physical measurements such as height, weight and blood pressure were collected in Step 2. Biochemical measurements were collected to assess blood glucose and cholesterol levels in Step 3. The STEPS survey was a population-based survey of adults aged 18-64. A multi-stage, cluster sample design was used to produce representative data for that age range in Samoa. A total of 1766 adults participated in the survey. The overall response rate was 64%. Age range of participants included: 18 to 64 years
*Source: Samoa STEPS Survey 2013 Fact Sheet. Available at: https://extranet.who.int/ncdsmicrodata/index.php/catalog/707*

**Seychelles: Seychelles Heart Study IV (Seychelles NCD Survey 2013)**
The survey was performed in a sex and age stratified random sample of all adults aged 25-64 years of Seychelles between October and December 2013 on Mahé and during 2 weeks in February 2014 in the islands of Praslin and La Digue. These three islands account for >98% of the total population of Seychelles. The eligible sample was extracted from the population registry. The survey was attended by 1240 adults, with a participation rate of 73%. Participants were invited to attend the survey on selected days in study centers located in Mahé, Praslin, and La Digue. All the eligible participants who did not attend were actively traced using (telephone, local administration, announcements on radio, etc) and invited to attend the survey. Since participants were randomly selected from the general adult population, findings of the survey can be inferred to the general adult population of Seychelles.
*Source: National Survey of Noncommunicable Diseases in Seychelles 2013-2014 (Seychelles Heart Study IV): methods and main findings. Available at: http://www.who.int/chp/steps/seychelles/en/.*

**Solomon Islands: STEPS 2015**
A multi-stage cluster sample design was used to produce representative data. Analysis weights were calculated by taking the inverse of the probability of selection of each participant. These weights were adjusted for differences in the age-sex composition of the sample population as compared to the target population.
Different weight variables are available per Step:
wStep1 - for interview data
wStep2 - for physical measures
wStep3 - for biochemical measures
This allows for differences in the weight calculation for each Step of the survey as the age-sex composition of the respondents to each Step can differ slightly due to refusal or drop out.
Age range of participants included: 18 to 69 years
*Source: no report or fact sheet available. Sampling information obtained from: https://extranet.who.int/ncdsmicrodata/index.php/catalog/710/study-description#page=overview&tab=study-desc*

**South Africa: SANHANES 2012**
The survey applied a multi-stage disproportionate, stratified cluster sampling approach. A total of 1000 census enumeration areas (EAs) from the 2001 population census were selected from a database of 86,000 EAs and mapped in 2007 using aerial photography to create the 2007 HSRC master sample to use as a basis for sampling of households. The selection of EAs was stratified by province and locality type. In the formal urban areas, race was also used as a third stratification variable (based on the predominant race group in the selected EA at the time of the 2001 census). The allocation of EAs to different stratification categories was disproportionate, in other words, over-sampling or over-allocation of EAs occurred in areas that were dominated by Indian, coloured or white race groups to ensure that the minimum required sample size in those smaller race groups were obtained. Based on the HSRC 2007 Master Sample, 500 Enumerator Areas (EAs) representative of the sociodemographic profile

of South Africa were identified and a random sample of 20 visiting points (VPs) were randomly selected from each EA, yielding an overall sample of 10 000 VPs. EAs were sampled with probability proportional to the size of the EA using the 2001 census estimate of the number of VPs in the EA database as a measure of size (MOS). One of the tasks of SANHANES-1 was to recruit and establish a cohort of 5 000 households to be followed up over the coming years. The sampling consisted of: Multi-stage disproportionate, stratified cluster sampling approach; 500 EAs within which 20 VPs/households per EA were sampled; Main reporting domains: sex (male, female), age-group (< 2 years, 2–5 years, 6–14 years, 15–24 years, 25–49 years, 50 years and older), race group (black African, white, coloured, Indian), locality type (urban formal, urban informal, rural formal [including commercial farms] and rural informal], and province (Western Cape, Eastern Cape, Northern Cape, Free State, KwaZulu-Natal, North West, Gauteng, Mpumalanga, Limpopo).
*Source: Human Sciences Research Council. SANHANES: Health and Nutrition. 2015. Available at: http://www.hsrc.ac.za/en/research-areas/Research_Areas_PHHSI/sanhanes-health-and-nutrition*

**Sri Lanka STEPS 2014**
A multi stage cluster sampling method was used to select a nationally representative sample from the total population. Department of Census and Statistics of Sri Lanka performed the selection of the study sample. Population of each divisional secretariat (DS) divisions as per the preliminary results of the Census done in 2012 was used for sampling.Sri Lanka is administratively divided in to 9 provinces and 25 districts. Each district is divided to Divisional Secretariat (DS) areas. Each DS area is divided to many Census Blocks, and each Census Block consists of many households.
The primary sampling unit (PSU) was a Divisional Secretariat (DS) area. Out of 331 DS areas available, 80 DS divisions were selected using proportionate to the size (PPS) sampling.A census block was considered as a SSU. From each DS division (PSU), six secondary sampling units (SSU) were selected using the proportionate to the size (PPS) sampling technique. Therefore, a total of 480 SSUs or census blocks were selected from 80 PSUs.

Number of houses in each census block depends on the area density and the population density in each DS division. Tertiary sampling unit (TSU) was the household and 15 households from each CB by random systematic sampling by the Department Census and Statistics. Therefore, a sample of 7200 (80x6x15) households were selected. In some instances, there were more than one household living in one house. People who are cooking and eating together were considered as one household. Whenever there were more than one household in a house, one household was selected randomly to be included in the study.
Only one participant from each household was included in the survey. All the eligible members in the selected family were listed in descending order according to the age. Once this was done, these data was fed to the personal digital assistants (PDAs). The PDAs then automatically selected the eligible participant using the Kish method.
Age range of participants included: 18 to 69 years
*Source: Sri Lanka STEPS 2014 Report. Available at: extranet.who.int/ncdsmicrodata/index.php/catalog/614/study-description#page=overview&tab=study-desc*

**St. Vincent & the Grenadines: STEPS 2013**

The survey covered the entire island St. Vincent and the Grenadines, and was conducted using the following zoning categories:

1) Mainland (St. Vincent)
2) Northern Grenadines (Bequia and Mustique)
3) Southern Grenadines (Canouan and Union Island)

The sample size was proportionately divided between the three main reporting strata (St.Vincent/Northern Grenadines/Southern Grenadines). The country's most recent age breakdown based on the 2001 national census by St. Vincent was used to approximate the adult population 18-69 years by Island grouping. The survey was stratified by sex, age groups 18-29, 30-44 and 45-69 years and by geographical location – St. Vincent, Northern Grenadines and Southern Grenadines.

A three-stage cluster sampling approach was used. Enumeration districts were randomly selected using Probability Proportional to Size (PPS) from the sampling frame. A total of 199 enumeration districts were selected. The sampling frame was developed using the number of households per enumeration district taken from the 2012 preliminary census report; enumeration districts had been subsequently revised (2010-2011) so that no enumeration district containing more than 150 Households would be randomly selected from the selected enumeration districts. The number of households per enumeration district to be selected was 26. Where an enumeration district had been split into 2 or more new enumeration districts the number of households in the previously defined enumeration district was divided equally between the newly revised enumeration districts. The household list for each selected enumeration district was updated prior to selection of households during a re-listing exercise. This was necessary as the existing household listing for each enumeration district was outdated.

Eligible persons at the household level were randomly selected using the Kish method. If no one was present in the selected household, a notification of visit card was left and the interviewer revisited. There was a total of three visits to the household before it was listed as non-response (one initial recruitment visit and two call backs). The interviewer then moved on to the next house on the list in the original order. Although the person selected for interview were to be at least 18 years and not older than 69 years on the last birthday, there were a few instances where some participants were turning 18 or 70 years; those cases were addressed during data cleaning.

Biological samples, testing and Nutrition intake (24 hour recall):

Fifty percent (50%) of the survey participants were asked to provide a biological specimen (finger prick) for Glucose and cholesterol testing using Glucose and Lipid Sampling Kits and respond to the nutrition intake (24 hour recall). The biological sample was only collected with participants' explicit consent; the samples were not stored or used for additional undetermined or undisclosed future testing to which respondents did not agree at the time of participation.

*Source: WHO STEPS: Noncommunicable Disease Risk Factor Surveillance. Report for St. Vincent & the Grenadines 2015. Available at: http://www.who.int/ncds/surveillance/steps/stvincent/en/*

**Sudan: STEPS 2015**

A four-stage cluster sampling design was implemented. The four sampling stages were; 1) selection of states from the six regions 2) selection of clusters (a cluster was a Popular Administrative unit), 3) selection of households and 4) selection of eligible individuals. First Stage (State): Administratively Sudan is divided into 18 states which are grouped in six regions, (North, East, Khartoum, Central, Kordofan and Darfur region (Table 1). States were randomly selected from each region. No geographical areas or populations were excluded from the sampling frame. Thus 11 states were selected, probability proportional to the size, to represent the six regions. A list of the selected states is shown in Table 2.1. Second Stage (Cluster PAU): The Popular Administrative Units (PAU) is the smallest geographically border unit. These were defined as the 'cluster' in the region. Clusters were randomly sampled from all PAUs, from both urban and rural strata, according to probability proportional to size in each state, and urban/rural distribution. The PAUs inaccessible due to security conditions were not excluded from the sampling frame, because within certain areas the security status was continuously changing. However, it was planned that if a PAU was found to be inaccessible at survey time, it should be replaced. However, no replacement was required during this survey. Third Stage (Household): Within the selected PAUs, all households (HH) were included in the sampling frame. Accordingly (HH) were selected using systematic random methods.

Fourth Stage (Individual): The members of the household were first listed in the mobile application (customized software). The inclusion criteria for the listed members were: all individuals aged between 18 to 69 years, from both sexes, irrespective of his health status and living in the selected household for a minimum of 6 weeks. The application was then run and it randomly selected the individual who will be selected to participate in the study. Age of participants included: 18-69 years.

*Source: Sudan STEPS 2015 report. Available at: https://www.who.int/ncds/surveillance/steps/Sudan_STEPwise_SURVEY_final_2016.pdf?ua=1*

**Tajikistan: STEPS 2016**

A multi-stage cluster sample of households. One individual within the age range of the survey was selected per household.

Analysis weights were calculated by taking the inverse of the probability of selection of each participant. These weights were adjusted for differences in the age-sex composition of the sample population as compared to the target population.

Different weight variables are available per Step:

wStep1 - for interview data

wStep2 - for physical measures

wStep3 - for biochemical measures

This allows for differences in the weight calculation for each Step of the survey as the age-sex composition of the respondents to each Step can differ slightly due to refusal or drop out. Age range of participants included: 18-69 years

*Source: report not available. Sampling information obtained from: extranet.who.int/ncdsmicrodata/index.php/catalog/270/study-description#page=sampling&tab=study-desc*

**Tanzania: STEPS 2012**

The STEPS survey in the United Republic of Tanzania was a population-based survey of adults aged 25-64. The study used both multistage cluster and random probability sampling procedures. Fifty of 119 total districts were randomly selected as primary sampling units (PSUs). Within these PSUs, enumeration areas (EAs) of > 50 households were randomly selected. Any EA with < 50 households was merged with a neighboring EA. Within the EAs, households were randomly selected from a list of all eligible households in the EA. A total of 5762 adults participated in the Tanzania STEPS survey. Within each selected household, the Kish method was used to select the STEPS participant. This procedure was followed until the predetermined sample was obtained for the enumeration area. The response rate for this survey was 94.7%.

*Source: Tanzania STEPS Survey Report. Available at: http://www.who.int/chp/steps/UR_Tanzania_2012_STEPS_Report.pdf?ua=1*

**Timor-Leste: STEPS 2014**

Note: Data from Census 2010 were used for all sampling considerations. Even though planning and mapping for 2015 Census is ongoing, data from the Census will only be available after July 2015.

STEP 1: Selection of Enumeration Area

(1) List of EA with number of HH by district for Census 2010 was obtained from the Directorate of Statistics. There are 1826 EAs in Timor-Leste. Out of these, 150 EAs were selected.

(2) The number of EAs to be selected from each district was based on their proportion in the country's population as per Census 2010.

(3) The numbers of Households (HH) per EAs varied from 0 to more than 300. Therefore, probability proportion to size (PPS) was used.

(4) For each district, the EAs were arranged in ascending order of HH size.

(5) Sampling interval was obtained by dividing the total number of HH in the district by the number of EA to be selected from that district.

(6) A random number was generated between one and the sampling interval for that district, using tools available at random.org.

(7) The EA where that random number fell was the first EA to be selected.

(8) Subsequently, the sampling interval was added to the random number and the EA where this new number fell was selected. For the next number, the sampling interval was added to the number and so on, till the population of HH was exhausted or target number of EA achieved.

(9) This was done separately for each district.

(10) The final list was compiled and had 150 EAs. These are spread over about 125 sucos.

STEP 2. Selection of Households in an Enumeration Area

Listing the house numbers to be visited

(1) It was decided to use the 2010 HH size of each EA. Based on past experience, it was expected that the increase would be on an average about 4–5%.

(2) The list of households to be selected by enumerators was decided centrally.
(3) Sampling interval was calculated by dividing the total number of households in the EA by 18.
(4) The first HH number was selected randomly by reading the last two digits of a currency note. If the number represented by the two digits was more than 18, the last digit was taken into consideration. For each EA, a different currency note was used. This could also be done it by using the tool at random.org. or by draw of lots.
(5) The subsequent HH are identified by adding the sampling interval as was done for selection of EA.
*Source: Timor-Leste STEPS Survey Report at http://www.who.int/entity/chp/steps/Timor-Leste_2014_STEPS_Report.pdf?ua=1*

**Togo: STEPS 2010**
Those included in this survey are male or female subjects, living in urban or rural areas, aged 15 to 64 on the day of the survey, residing in the enumeration area for at least 6 months and having given their informed consent to participate in this study. [...] Three hundred clusters were randomly selected in a systematic draw with probability proportional to the size of the cluster (number of households) in the 4620 areas of enumeration of the DGSCN (General Directorate of Statistics and National Accounts) sampling frame. In order to obtain the 4,800 households at the rate of 1 individual / household, 16 households per cluster were randomly selected at the second stage of survey. In each of the selected households, one individual was selected as a survey participant via the Kish Method. A household was defined as the group of persons, who regularly share the main meal (regardless of their relationship). Households were not replaced in the event of a refusal or two unsuccessful visits to the eligible person selected by Kish's method. If the selected person was unwell or not present at the time of the interview, the investigators either tried to find a new appointment or searched for the respondent.
*Source: Translated from WHO: The Final Report on the Togo STEPS Survey 2010. Available at: http://www.who.int/chp/steps/2010STEPS_Report_Togo_FR.pdf?ua=1.*

**Tokelau STEPS 2014**
A whole population-based (census) survey was used to produce representative data for that age range in Tokelau. Analysis weights contain adjustments for differences in the age-sex composition of the sample population as compared to the target population. Different weight variables are available per Step:
wStep1 - for interview data
wStep2 - for physical measures
wStep3 - for biochemical measures
This allows for differences in the weight calculation for each Step of the survey as the age-sex composition of the respondents to each Step can differ slightly due to refusal or drop out. *Source: Report unavailable. Sampling information obtained from: https://extranet.who.int/ncdsmicrodata/index.php/catalog/638/overview#page=sampling&tab=stu dy-desc*

**Tonga: STEPS 2017**
An initial sample of 4,500 individuals (respondents) between the ages of 18 to 69 years old was targeted to undertake the STEPS survey for 2017 in Tonga.

Because it is important to compare the results by island divisions (national level), it is required with importance to produce the estimates in the divisional level (National Level). Therefore the sampling fractions will be adjusted from its proportional to the size (number of households) to have higher sampling fraction (coverage) for the smaller size island division as shown in the following table:
Pop Census STEPS sample

*Appendix table A2.5: Tonga island survey coverage and numbers of blocks selected*

| | Island | Total HH | Ideal sample size | Coverage | Number of selected blocks |
|---|---|---|---|---|---|
| 1 | Tongatapu | 12953 | 3240 | 25.0% | 270 |
| 2 | Vava'u | 2,715 | 684 | 25.2% | 57 |
| 3 | Ha'apai | 1,179 | 288 | 24.4% | 24 |
| 4 | Eua | 885 | 228 | 25.8% | 19 |
| 5 | Niua | 273 | 60 | 22.0% | 5 |
| | **Total** | **18,005** | **4,500** | **25.0%** | **375** |

The final sample numbers presented in the table above were rounded such that they were divisible by 12 (an enumerators workload) to accommodate field logistics. As such the sample size is recorded to 4,500. The sample was selected independently within each of the 5 target areas.

The sampling in each area was then undertaken using a three-stage process. The first stage involved the selection of census blocks using Probability Proportional to Size (PPS) sampling, where the size measure was the expected number of households in that block. For the second stage, a fixed number (twelve) of households were selected from each selected census block using systematic sampling. The household lists for all selected blocks were updated just prior to the second stage of selection. Once the selected 12 households are found, then the list of household members age 15 to 64 by gender will be recorded. The final stage will be to use the Random Sample Generator (Android Application) to randomly select one person from the household to be enumerated so that it captures the required composition of the sample with specific age-group distribution and gender.
Age range of participants included: 18 to 69 years

*Source: Tonga STEPS Survey 2017 Sampling Design. Available at: https://pacificdata.org/data/dataset/spc_ton_2017_steps_v02_m/resource/261f0a3c-4979-4a42-a560-1b103c617a42?inner_span=True*

**Tuvalu: STEPS 2015**
The Tuvalu STEPS Survey was a population based cross-sectional survey of 18-69 year olds. Analysis weights were calculated by taking the inverse of the probability of selection of each

participant. These weights were adjusted for differences in the age-sex composition of the sample population as compared to the target population.
Different weight variables are available per Step:
wStep1 - for interview data
wStep2 - for physical measures
wStep3 - for biochemical measures
This allows for differences in the weight calculation for each Step of the survey as the age-sex composition of the respondents to each Step can differ slightly due to refusal or drop out. Additionally, some countries perform subsampling for Step 2 and/or Step 3. When no subsampling is done and response rates do not differ across Steps of the survey, the 3 weight variables will be the same.
Age range of participants included: 18 to 69 years
*Source: no report or fact sheet available. Sampling information obtained from: https://extranet.who.int/ncdsmicrodata/index.php/catalog/639/study-description#page=overview&tab=study-desc*

**Uganda STEPS 2014**
Uganda has a total population of 34.9 million people, approximately 43% of which are adults aged 18 years or older [14]. The survey covered the whole country, and a three stage sampling design was used to select participants. The sampling procedure utilized the Uganda Bureau of Statistics (UBOS) master sampling frame of Enumeration Areas (EAs) that had just been demarcated throughout the country in preparation for the 2014 population and housing census. Each EA included 150–200 households. In the first stage, a random sample of 350 out of 78,950 EAs was selected with selection probability proportional to the size (PPS) of the number of households in the EAs. The EAs were stratified across the four regions of Uganda namely: Central, Eastern, Northern and Western region; and were selected with separate estimates for rural and urban areas. Urban areas were defined as EAs within government designated urban areas, or those within other geographic divisions with population density of more than 1000 per square kilometer.

After selecting the 350 EAs, trained teams of UBOS staff were dispatched throughout the country to list the households and their household heads within the 350 EAs. A household was defined as a group of individuals that usually shared meals together, and had a household head who usually made major decisions for the household. In the second stage of sampling, 14 households were randomly selected from the listed households in each of the sampled EAs.

Research Assistants (RA) that had received a five-day training on procedures and administration of the STEPs tool, enumerated eligible household members who were recorded in Personal Digital Assistants (PDA), which was then used to randomly select one subject for inclusion in the survey giving a total sample of 4900. Eligible subjects were household members aged 18 to 69 years, who had resided in the sampled households for at least six months preceding the date of interview.
*Source: Guwatudde D, Mutungi G, Wesonga R, Kajjura R, Kasule H, Muwonge J, et al. (2015) The Epidemiology of Hypertension in Uganda: Findings from the National Non-Communicable Diseases Risk Factor Survey. PLoS ONE 10(9): e0138991. doi:10.1371/journal.pone.0138991.*

**Vanuatu: STEPS 2011**

The survey used a cluster sampling design where the primary sampling unit was enumeration area (EA) and the secondary sampling unit was households. All 6 provinces in Vanuatu were included in the survey. One hundred and thirteen (113) EAs were randomly selected proportion to the size of the EA from a total of 411 EAs. Forty four (44) households were then randomly selected in each EA proportional to the number of households in each EA. The selection of participants within each household was done using the Kish method. The total number of households selected by combined Enrolment Areas was 4,972.

The required sample size was calculated as 4972 households on a margin of error of 0.05, an anticipated response rate of 89% and with 80% power to detect statistically significant differences between six age/sex groups. Accordingly, from the 4,972 selected households 4,649 individuals aged 25-64 years participated in STEP 1 and STEP 2 giving an overall response rate of 94%. The response rate dropped to 85% for STEP 3 with 4,224 people participating.

Age range of participants included: 25 to 64 years

*Source: Vanuatu STEPS report [online]: https://extranet.who.int/ncdsmicrodata/index.php/catalog/714*

**Vietnam: STEPS 2015**

At the same time of STEP survey, MOH also conduct the Global Adult Tobacco Survey (GATS) at the same scale, location, and study subjects (>15 years for GATS and 18-69 for STEPS). The sampling of STEPS was done in as part of the sampling for the (GATS) conducted in combination manner to save time and resources for these two surveys. Applied the multi-stages complex sampling process, the sampling process done by GSO was as follow: • Sampling of clusters (EA) In the first stage of sampling, the primary sampling unit (PSU) was an enumeration area (EA). There are about 170,000 EAs in the whole Viet Nam and the average number of households in each EA is different between urban and rural areas. An average number of households in an urban EA and a rural EA is 133 households and 120 households, respectively. Sample of EAs were selected from the master sample frame. The master sample frame was a cluster frame made by the GSO based on the frame of Population and Housing Census 2009 and updated with data of 2014. Based on the Population and Housing Census data 2009, GSO prepared a 15% of master sample to serve as a national survey sampling frame. The master sample frame contains 25,500 enumeration areas (EAs) from 706/708 districts of Viet Nam (2 island districts were excluded from the GSO master sample frame). The master sample frame of GSO was divided by two stratification variables: urbanization (1 = urban; 2 = rural) and district group (1 = district/town/city of province; 2 = plain and coastal district; 3 = mountainous, island district). It means that the master sample frame was divided into 6 sample frames or 6 strata. The probability proportional to size (PPS) sampling method was used to select sample of EAs from 6 strata of master sample frame. The final sample of GATS included 315 EAs in the urban and 342 EAs for the rural. From these 657 EAs, 315 EAs were systematically selected for STEPS.

Sampling of households At the second stage of sampling, 10% households in each EA were selected. Thus, 15 households from the selected urban EA and 14 households from the selected rural EA were chosen using simple systematic random sampling. The total households for STEPS 2015 were 4,651 households.

Sampling of individuals: One eligible person is then randomly selected from each selected household for the STEPS 1 interview. The selection of individual is automatically done by the PDA program after eligible household members are entered into the PDA. The selection probability of an eligible individual was calculated as a product of selection probability for each stage. The sampling base weight for an eligible individual was the inverse of the selection probability shown above.
Age range of participants included: 18 to 69 years
*Source: National Survey on the Risk Factors of Non-communicable diseases (STEPS) Viet Nam Report 2015. Available at: https://www.who.int/ncds/surveillance/steps/viet_nam/en/*

**Zambia: STEPS 2017**
To ensure that the sample reflected the entire country of Zambia, a multi-stage cluster sampling technique was used to select a nationally representative sample of adults in Zambia aged 18 to 69 years. It was decided to utilize the household listing from the Zambia PopulationBased HIV Impact Assessment (ZAMPHIA) - a household-based national survey that was conducted between March and August 2016 in order to measure the status of Zambia's national HIV response. ZAMPHIA offered the most pragmatic up to date and accessible national household listing to be used as the sampling frame for this survey. The ZAMPHIA survey included 60,581 households drawn from 1,103 clusters referred to in this report as standard enumeration area (SEA) (Table 2.4.1). Thus the sample drawn for the STEPS survey was a subsample of the households selected for the ZAMPHIA survey. In the first stage of sampling, SEAs were selected from each province using probability proportional to size (PPS). In the second stage, 15 households in rural SEAs and 20 households in urban SEAs were selected systematically using appropriate sampling interval based on the number of households in that SEA. These households constituted the final list of households for the STEPS survey prepared for the field investigators (FI). In the third stage, while the FI approached the household and sought consent, all eligible members in the household were entered into the Android-based devise used for the survey. The device then selected one member from the eligible members using a simple random sampling technique. The selected member was then interviewed having gone through the ethical process of consent after being provided with information on the survey. If the selected member was not available, a scheduled visit was made. If the selected member could not be reached after two scheduled visits he or she was considered as non- response. There was no replacement strategy so as to maintain the integrity and representativeness of the sample.
Age range of participants included: 18 to 69 years
*Source: STEPS 2017 Report. Available at: https://extranet.who.int/ncdsmicrodata/index.php/catalog/620*

**Zanzibar: STEPS 2011**
The survey took place in June and July 2011, followed by data cleaning and analysis. One Principal Investigator and five assistant researchers coordinated the survey on site, checked completed questionnaires daily, and organized logistics. The six data collection teams consisted each of six interviewers, one supervisor, one laboratory technician and one driver. Interviewers were either health care workers or professional interviewers familiar with household surveys such as DHS. The sample size was calculated to be 2800 participants. Each

interviewer did on average 3 – 4 interviews a day and was assisted on site by local village guides.

The study was a cross-sectional population based survey with a sample of a sufficient size with a power to determine the proportion of adults that are exposed to selected risk factors associated with NCDs; including those having raised BP, FBG or blood lipids, had experienced injuries or traumas in recent times, and/or were mentally unwell (anxiety, depression), as well as linking these conditions with one another and with the sociodemographic and economic information obtained. People reported to be permanent residents (spending on average maximum 3 nights per week outside the house, and not holding an address in another place) in the selected households and fulfilled the inclusion criteria were enrolled into the survey. A person could only appear once in the study. Therefore we classified a husband practicing polygamy to be listed in the household of his first wife but not to be a member in the household of the following wives. Inclusion criteria was age between 25 - 64 years, able to understand the information given by the interviewer about the study prior to the beginning of the interview, signing of the informed consent for accepting participation. Exclusion criteria was inability to understand or comprehend the information given by data collector, inability to communicate through verbal expression for consent and for responding to the questionnaires, severe/terminal illness that hinders participation in the survey.

The target population is the entire population in Zanzibar whereby the whole of Zanzibar was selected as the survey site, and hence all districts included. The total population is estimated to be 1.2 million distributed unevenly between 10 districts. The sampling frame represented the entire population in Zanzibar. The sampling strategy used is a multi-stage cluster sampling with stratification. The ten districts are considered as different strata, and the total number of primary sampling units, PSU, is allocated proportionately across all strata. Each district is divided into smaller clusters. These clusters are the geographical and administrative units called Shehia11. The Shehia are divided into smaller clusters called zones (also called mitaa, vitongoji, or vijiji) which typically consist of 100-300 households. Zones smaller than that were merged to make up one larger cluster, and zones much larger were split in smaller clusters.

At the first stage clusters were selected using Simple Random Selection, SRS, from the list of clusters (Shehia) within each district. At the second stage clusters (zones) were randomly selected using probability proportionate to size (PPS). At the third stage households were randomly selected from the household lists provided by the administrative leader of the Shehia. The two last stages of sampling were done using the software STEPSsampling.xls from WHO. Finally participants were selected from the household using Kish method. The household lists were complete and included households with no eligible participants for the survey. Therefore an extra 7 households were sampled at third stage in each cluster for replacement in case a selected household had no eligible participants and had to be changed. This was done before data collectors went to the cluster.

Resources allowed for 100 PSU which was why 2800/100 = 28 households were selected from each PSU (and disproportionate from each SSU). A structured questionnaire was used, based on WHO STEPwise approach to chronic diseases risk factor surveillance.. After getting behavioural and socio-demographic information, anthropometric measurements (BP, height, weight, waist and hip circumference) was done the same day. Answers were recorded electronically during interview using a Personal Digital Assistant (PDA). Biochemical measurements (fasting blood glucose, triglyceride, and cholesterol levels) were done the next day at a central place in each study site according to appointment and were done by Laboratory technicians using dry chemistry for rapid and convenient results and to avoid

suspicion surrounding sending away blood samples. Results were recorded electronically on site using a PDA, and participants received a paper copy of the results.
Every study site was visited one day for interviews. Sampled households/ participants were visited at least three times before recorded as non-respondent. The following day the site was visited for biochemical measurements. Laboratory technicians called participants who did not show up to ask them to set up appointment for the following day (at a new study site). After all study sites had been visited call-backs were made to all eligible participants (non-respondents) who's number we had obtained. A time and place near the participants was identified for data collection. Participants met fasting and started with having blood sample drawn, afterwards the interviews and anthropometric measurements were conducted. Laboratory technicians continued biochemistry measurements for another few days.
Age range of participants included: 25 to 69 years
*Source: Zanzibar STEPS Survey Report, [online]*
*https://www.who.int/ncds/surveillance/steps/2011_Zanzibar_STEPS_Report.pdf*

## Appendix 2.6: Detailed methodology for household wealth index calculation

Across surveys, several different wealth indicators were measured including continuous income, income categories, income quintiles, an asset index, or a combination of these (see table below). In an effort to homogenize wealth in the pooled analysis, we constructed household wealth quintiles for each survey.

*Appendix table A2.6: Measures used for wealth index calculation by country*

| Wealth Measure | Country |
|---|---|
| Asset index | India, Indonesia, Iran, Kenya, Namibia |
| Continuous income | Bhutan, Brazil, The Gambia, Ecuador, Eritrea, Kiribati, Laos, Myanmar, Romania, Timor Leste, Tuvalu |
| Continuous income and quintiles | Algeria, Azerbaijan, Benin, Botswana, Comoros*, Eswatini*, Georgia, Kyrgyzstan, Lesotho, Liberia*, Moldova, Rwanda*, Samoa*, Solomon Islands, Tajikistan, Tanzania, Togo*, Uganda*, Vanuatu, Zambia, Zanzibar |
| Continuous income and categories | Benin, Guyana, Lebanon, Mongolia*, Morocco |
| Income categories only | Seychelles, Sri Lanka, South Africa, St. Vincent & the Grenadines, Sudan |
| No wealth indicators assessed | Bangladesh, Belarus, Burkina Faso, Cambodia, Costa Rica, Iraq, Nepal, Tokelau, Tonga, Vietnam |

*Quintiles were not used as they displayed large discrepancies with respect to continuous income range or could not be correctly identified

The construction of wealth quintiles depends on the given wealth indicator. Countries using an asset index surveyed a range of assets, dwelling characteristics, and further country-specific variables. Utilizing the standard DHS approach, we used principle component analysis to derive an asset index, from which we create unweighted wealth quintiles. Countries using an income-based measurement mainly followed the STEPS template questionnaire put forward by the WHO. In this, respondents were asked about the average earnings (taking the past year) of the household in a week, month, or year. In cases where this question was left unanswered, a pre-coded estimate of the households' annual income was indicated. This pre-coded estimate was usually expressed as quintiles and sometimes as categories that were defined by the countries' survey teams. Using both the pre-coded estimates as well as the continuous income, we again created unweighted wealth quintiles. In this, we assumed that national incomes follow a log-normal distribution and made use of the procedure put forward by Harttgen and Vollmer[2] in combining income quintiles and categories. In seven cases, we dismissed pre-coded quintiles or income as they displayed very large discrepancies with respect to the continuous income range or could otherwise not be correctly identified. However, as the pre-coded estimates were typically only asked of respondents that had not indicated a continuous income, this led to only minor information losses.

## Appendix 2.7: Logistic regression equations

For each CVD risk factor, the following multivariable logistic regression models are estimated:

$$Screened_{ij} = \alpha_0 + \beta_j + \alpha_1 Sex_{ij} + \alpha_2 Wealth_{ij} + \alpha_3 Education_{ij} + \epsilon_{ij}$$

Individual $i$
Country $j$
Screening status dummy (0 = not screened, 1 = screened) $Screened_{ij}$
Country-fixed effect $\beta_j$
Sex dummy (0 = male, 1 = female) $Sex_{ij}$
Wealth quintile $Wealth_{ij}$
Education category (Less than primary, Less than secondary, Secondary or more) $Education_{ij}$

We did not test for normality or homogeneity of variance as these assumptions are only relevant for small samples and each CVD risk factor sample is large. We also did not test for linearity of our quantitative predictors as each numerical explanatory variable was categorical.

## Appendix 2.8: Countries' WHO World Regions categories

*Appendix table A2.7: WHO World Regions categories*

| **WHO World Regions** | **Countries included in analysis of main article** |
|---|---|
| Africa | Algeria, Benin, Botswana, Burkina Faso, Comoros, Eritrea, Eswatini, Gambia, Kenya, Lesotho, Liberia, Namibia, Rwanda, Seychelles, South Africa, Tanzania, Togo, Uganda, Zambia, Zanzibar |
| Americas | Brazil, Costa Rica, Ecuador, Guyana, St. Vincent & the Grenadines |
| South East Asia | Bangladesh, Bhutan, India, Indonesia, Myanmar, Nepal, Sri Lanka, Timor-Leste |
| Western Pacific | Cambodia, Kiribati, Kyrgyzstan, Marshall Islands, Mongolia, Samoa, Solomon Islands, Tokelau, Tonga, Tuvalu, Vanuatu, Vietnam |
| Europe | Azerbaijan, Belarus, Georgia, Kyrgyzstan, Moldova, Romania, Tajikistan |
| Eastern Mediterranean | Iran, Iraq, Lebanon, Morocco, Sudan |

# Appendix 2.9: Main results – diagnostic testing performance by CVD risk factor

*Appendix table A2.8: Diagnostic testing performance by CVD risk factor*

| CVD Risk factor | Dimension | Overall (in %) |
|---|---|---|
| Hypertension | Fulfills diagnostic testing criteria | 19·1 (18·5-19·8) |
| | Tested | 63·1 (62·4-63·8) |
| | Tested out of all fulfilling criteria | 78·6 (77·8-79·2) |
| | Guidelines adhered | 49·0 (48·7-49·4) |
| Diabetes | Fulfills diagnostic testing criteria | 23·8 (23·4-24·3) |
| | Tested | 28·6 (28·0-29·2) |
| | Tested out of all fulfilling criteria | 44·9 (43·7-46·2) |
| | Guidelines adhered | 72·2 (71·7-72·7) |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 27·4 (26·3-28·6) |
| | Tested | 29·8 (26·9-32·9) |
| | Tested out of all fulfilling criteria | 39·7 (37·1-42·4) |
| | Guidelines adhered | 70·6 (69·7-71·4) |

| Sex | | Female (in %) | Male (in %) | p-value from two-sided t-test |
|---|---|---|---|---|
| Hypertension | Fulfills diagnostic testing criteria | 21·6 (21·0-22·2) | 16·5 (15·7-17·2) | <0·0001* |
| | Tested | 70·3 (69·7-70·9) | 56·3 (55·4-57·1) | <0·0001* |
| | Tested out of all fulfilling criteria | 82·0 (81·3-82·6) | 74·1 (72·9-75·2) | <0·0001* |
| | Guidelines adhered | 44·4 (44·1-44·8) | 53·3 (52·8-53·8) | <0·0001* |
| Diabetes | Fulfills diagnostic testing criteria | 28·9 (28·3-29·5) | 18·4 (17·8-18·9) | <0·0001* |
| | Tested | 31·5 (30·8-32·3) | 25·6 (24·8-26·4) | <0·0001* |
| | Tested out of all fulfilling criteria | 45·6 (44·1-47·1) | 44·6 (42·7-46·6) | 0·4595 |
| | Guidelines adhered | 69·4 (68·8-69·9) | 75·2 (74·5-75·9) | <0·0001* |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 32·4 (31·5-33·4) | 22·3 (20·8-24·0) | <0·0001* |
| | Tested | 31·8 (28·9-34·9) | 27·7 (24·9-30·8) | 0·0612 |
| | Tested out of all fulfilling criteria | 39·7 (37·0-42·5) | 40·2 (37·5-42·9) | 0·8119 |
| | Guidelines adhered | 67·4 (66·6-68·1) | 73·8 (72·6-74·9) | <0·0001* |

| Wealth | | Poorest quintile (in %) | Middle quintile (in %) | Richest quintile (in %) | p-value from two-sided t-test |
|---|---|---|---|---|---|
| Hypertension | Fulfills diagnostic testing criteria | 11·7 (11·1-12·3) | 18·0 (17·2-18·8) | 27·5 (26·7-28·3) | <0·0001* |
| | Tested | 48·8 (47·7-49·9) | 61·5 (60·6-62·5) | 73·2 (72·0-74·4) | <0·0001* |
| | Tested out of all fulfilling criteria | 62·0 (59·8-64·2) | 74·8 (73·3-76·1) | 84·2 (83·1-85·2) | <0·0001* |
| | Guidelines adhered | 57·3 (56·6-57·9) | 49·2 (48·6-49·8) | 46·2 (45·3-47·0) | <0·0001* |
| Diabetes | Fulfills diagnostic testing criteria | 22·4 (21·5-23·3) | 27·1 (26·2-28·2) | 33·5 (32·0-35·0) | <0·0001* |
| | Tested | 20·1 (19·2-21·0) | 26·9 (25·9-27·9) | 37·1 (35·3-39·0) | <0·0001* |
| | Tested out of all fulfilling criteria | 31·4 (29·0-33·9) | 40·7 (37·8-43·6) | 51·8 (49·6-54·0) | <0·0001* |
| | Guidelines adhered | 76·0 (75·2-76·9) | 72·0 (71·1-73·0) | 67·6 (65·9-69·3) | <0·0001* |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 24·9 (23·1-26·7) | 31·8 (29·9-33·7) | 37·6 (35·9-39·3) | <0·0001* |
| | Tested | 27·9 (23·6-32·6) | 35·7 (30·7-41·1) | 45·0 (40·4-49·6) | <0·0001* |

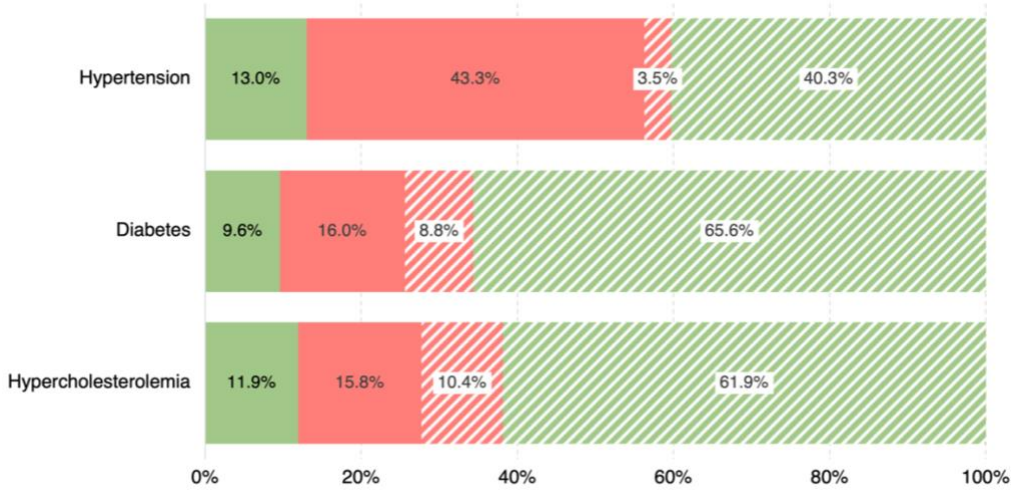
| | | | | | |
|---|---|---|---|---|---|
| | Tested out of all fulfilling criteria | 36·2 (32·1-40·5) | 44·3 (39·5-49·1) | 53·4 (49·5-57·1) | 0·0002* |
| | Guidelines adhered | 64·6 (62·2-67·0) | 64·1 (62·2-65·9) | 60·8 (58·3-63·2) | <0·0001* |
| **Education** | | **Less than primary school (in %)** | **Less than secondary school (in %)** | **Secondary school or more (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 20·5 (19·7-21·3) | 21·0 (20·4-21·6) | 20·3 (19·3-21·3) | 0·7374 |
| | Tested | 59·7 (58·7-60·7) | 62·0 (61·1-62·9) | 62·7 (61·7-63·7) | <0·0001* |
| | Tested out of all fulfilling criteria | 72·4 (70·7-74·0) | 75·5 (73·2-77·5) | 77·6 (75·8-79·4) | <0·0001* |
| | Guidelines adhered | 52·1 (51·3-52·8) | 51·3 (50·4-52·1) | 49·8 (49·0-50·7) | 0·0001* |
| Diabetes | Fulfills diagnostic testing criteria | 27·6 (26·4-28·9) | 26·2 (25·1-27·4) | 25·2 (24·1-26·3) | 0·0033* |
| | Tested | 25·2 (23·9-26·6) | 26·5 (25·3-27·8) | 27·1 (25·9-28·2) | 0·0452 |
| | Tested out of all fulfilling criteria | 38·0 (34·6-41·4) | 40·2 (37·2-43·3) | 44·2 (41·5-47·0) | 0·0049* |
| | Guidelines adhered | 74·4 (73·1-75·6) | 73·6 (72·4-74·8) | 73·5 (72·4-74·7) | 0·3249 |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 30·9 (28·4-33·6) | 32·6 (31·5-33·7) | 28·8 (27·1-30·5) | 0·1735 |
| | Tested | 25·9 (22·3-29·9) | 27·2 (25·0-29·5) | 27·3 (23·8-31·0) | 0·6098 |
| | Tested out of all fulfilling criteria | 30·8 (27·3-34·6) | 33·9 (31·4-36·6) | 37·1 (33·4-41·0) | 0·0189* |
| | Guidelines adhered | 71·8 (70·0-73·5) | 72·8 (71·8-73·7) | 72·3 (71·2-73·4) | 0·5964 |

Note: 95% confidence intervals in brackets. The p-values were calculated using an immediate form of a two-sample t-test with unequal variances. We compared male vs. female sex, less than primary vs. secondary or more education, and richest vs. remaining four wealth quintiles. Statistical significance levels based on Benjamini-Hochberg adjusted p-values corresponding to a 5% significance level represented with *.

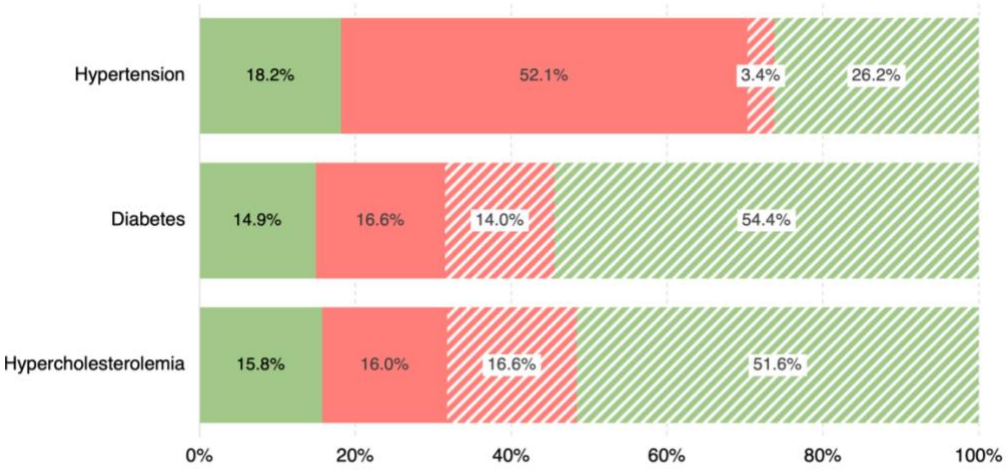
# Appendix 2.10: Main results – bar charts by sex, wealth, and education

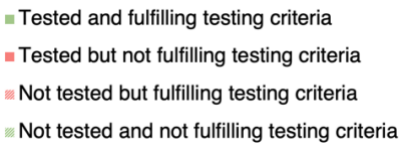
*Appendix figure A2.4: WHO PEN diagnostic testing recommendations and testing status by sex*

## Male



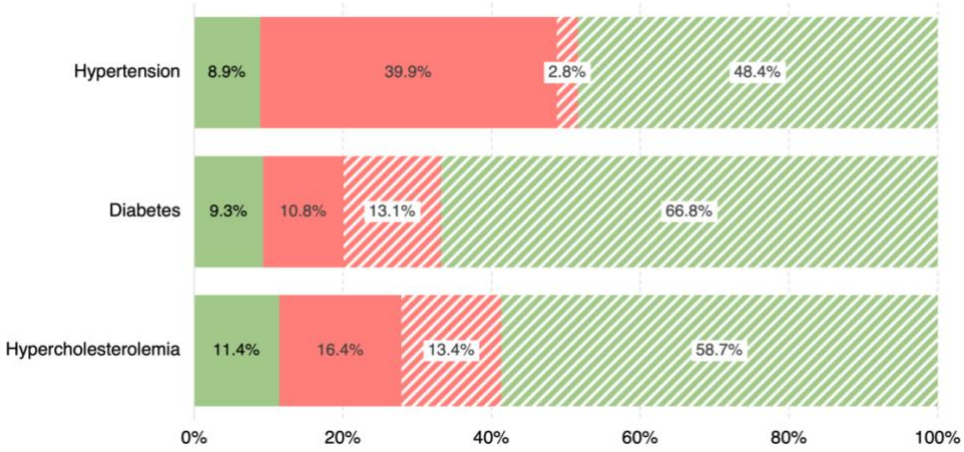

## Female



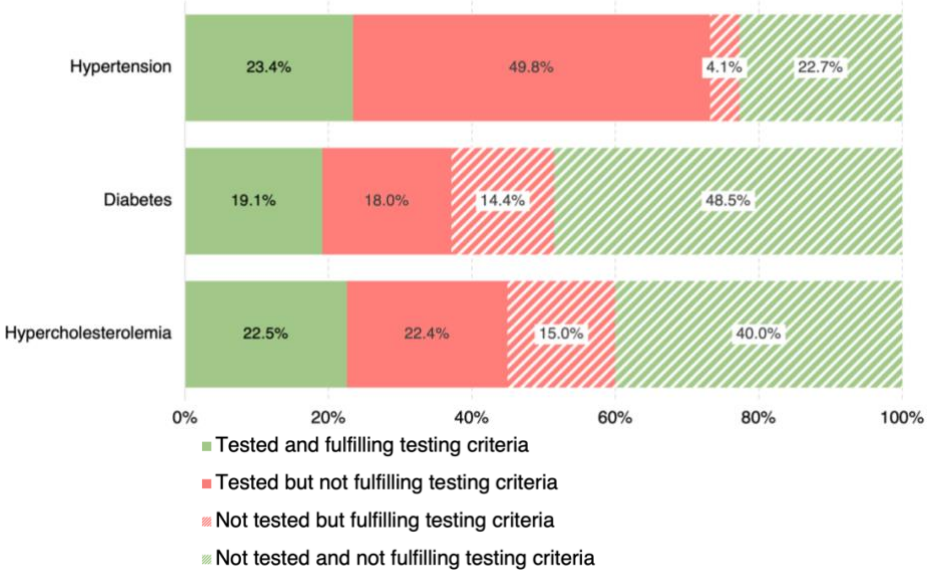
*Appendix figure A2.5: WHO PEN diagnostic testing recommendations and testing status by wealth quintile*

Poorest wealth quintile



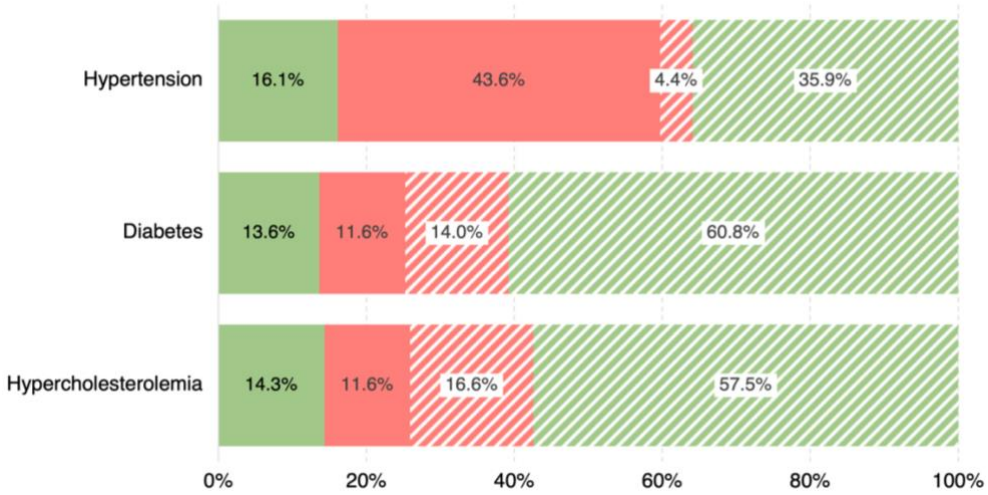

Middle wealth quintile


Richest wealth quintile



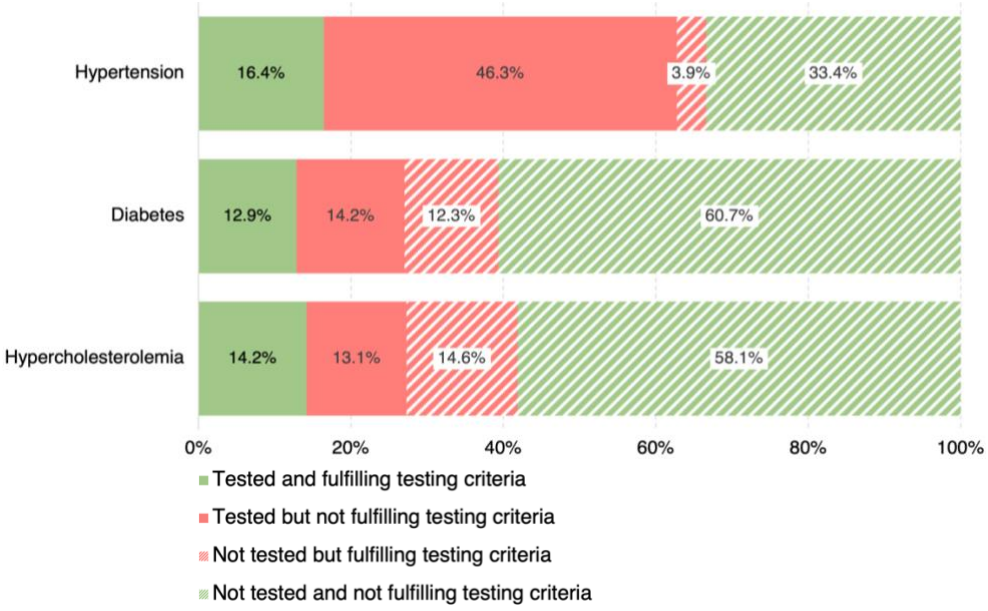
*Appendix figure A2.6: WHO PEN diagnostic testing recommendations and testing status by education category*

Less than primary education




Less than secondary education

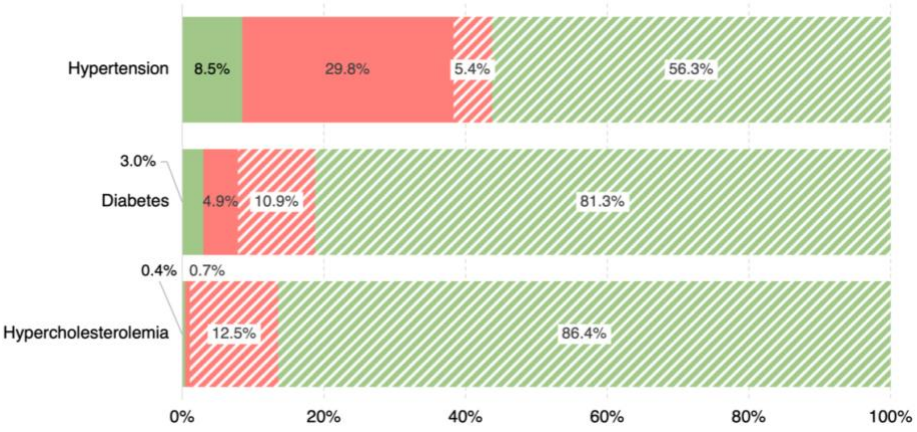

Secondary education or more



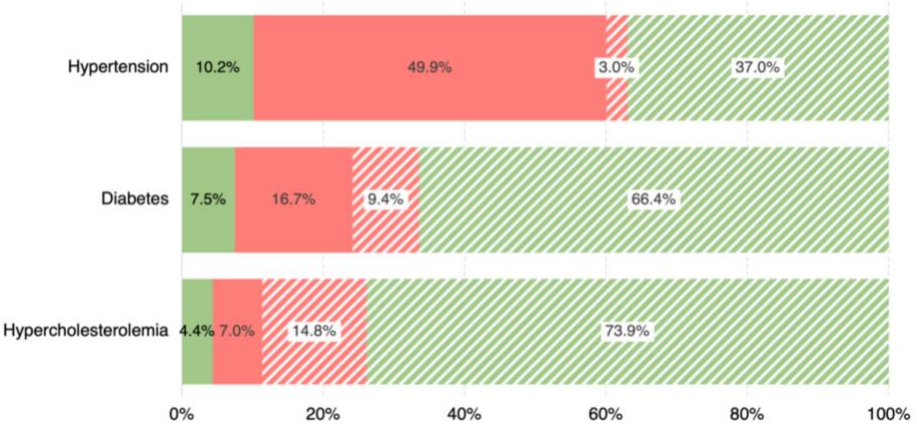
## Appendix 2.11: Main results – diagnostic testing performance by Income Group and World Region

*Appendix figure A2.7: WHO PEN diagnostic testing recommendations and testing status by World Bank Income Group*

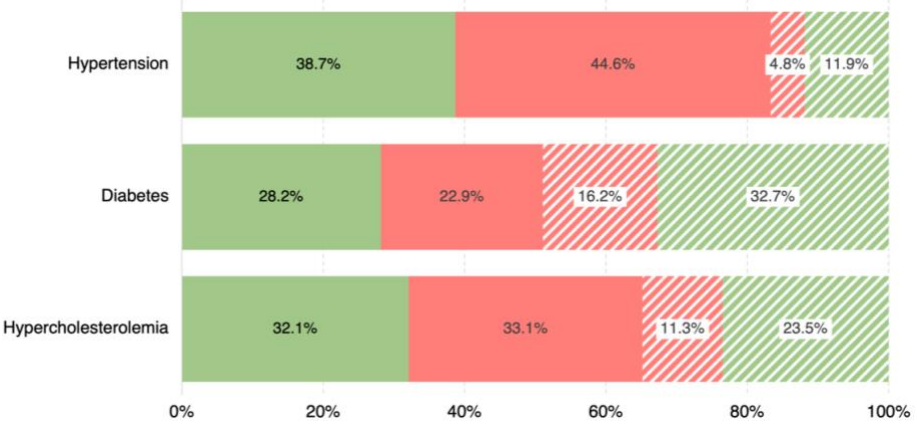
### Low-Income Countries




### Lower-Middle-Income Countries



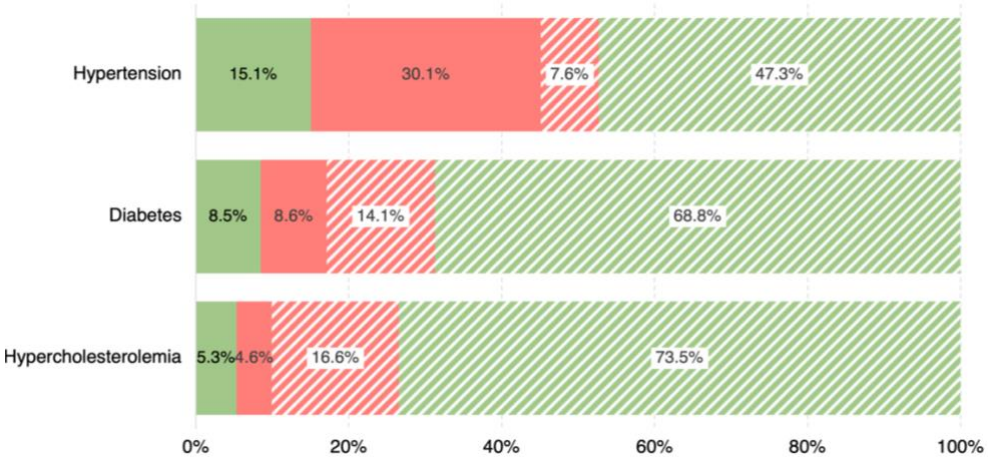

### Upper-Middle-Income Countries



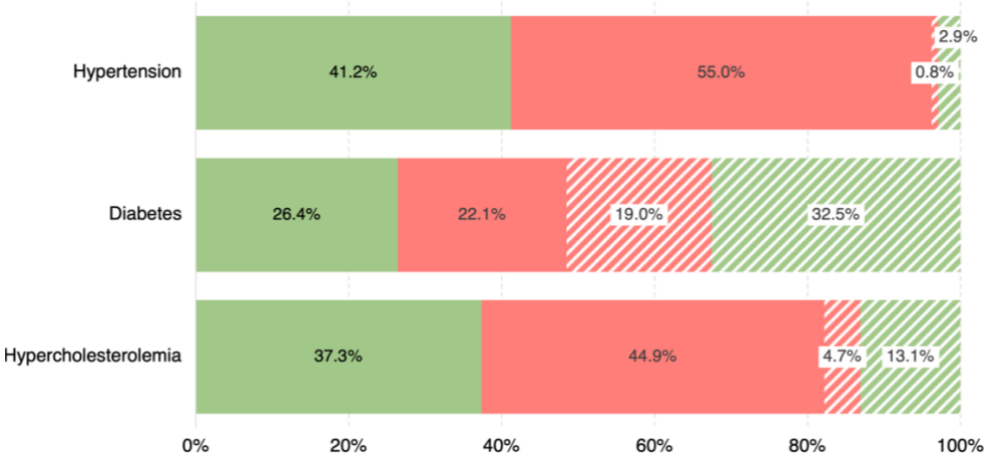

- Tested and fulfilling testing criteria
- Tested but not fulfilling testing criteria
- Not tested but fulfilling testing criteria
- Not tested and not fulfilling testing criteria

*Appendix figure A2.8: WHO PEN diagnostic testing recommendation and testing status by WHO World Region*

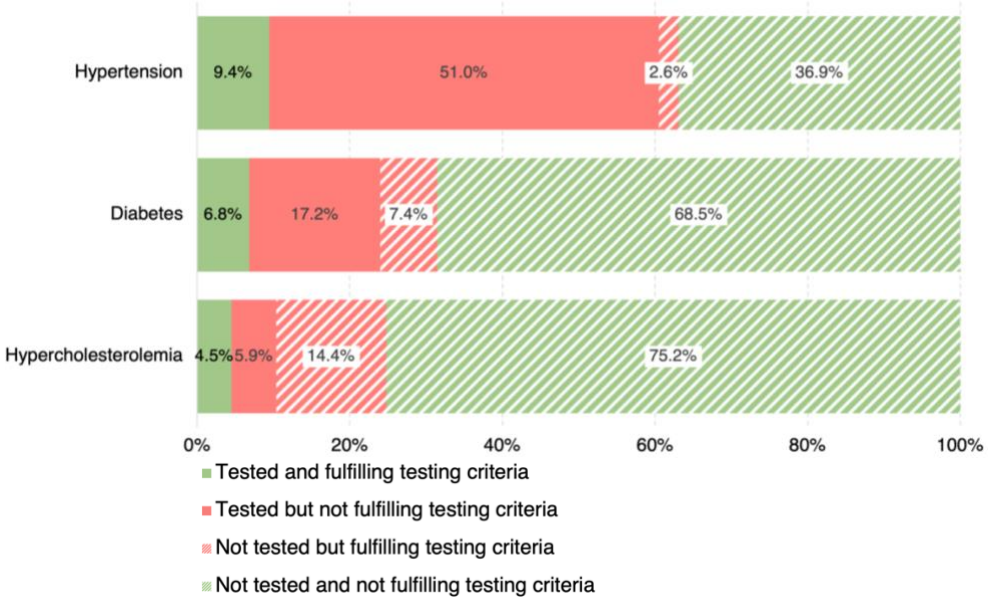
WHO World Region: Africa



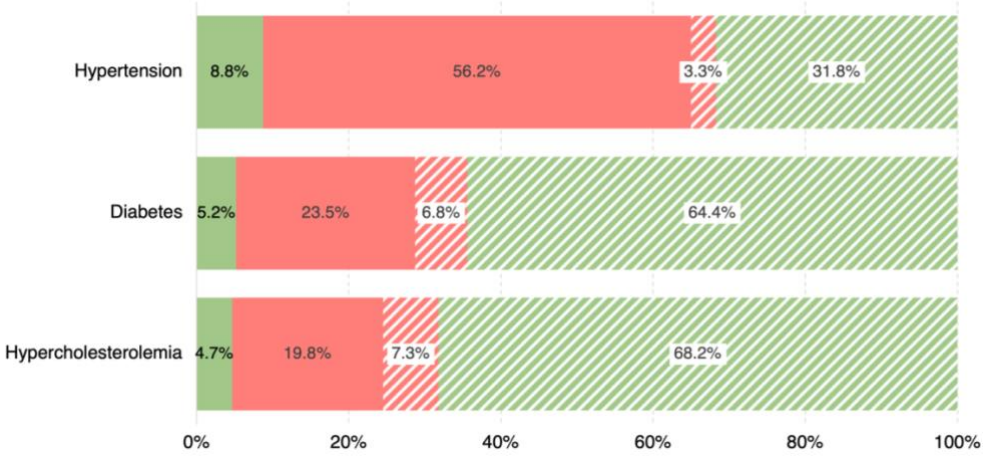

WHO World Region: Americas



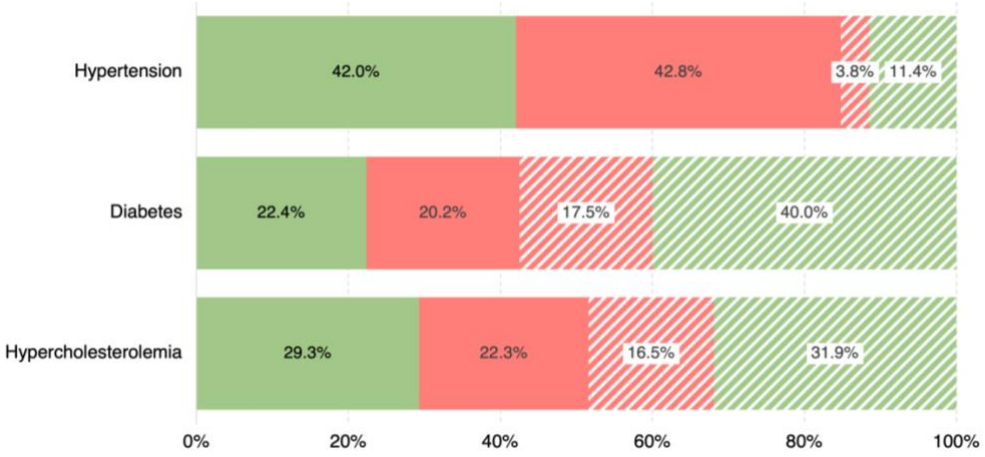

WHO World Region: South East Asia



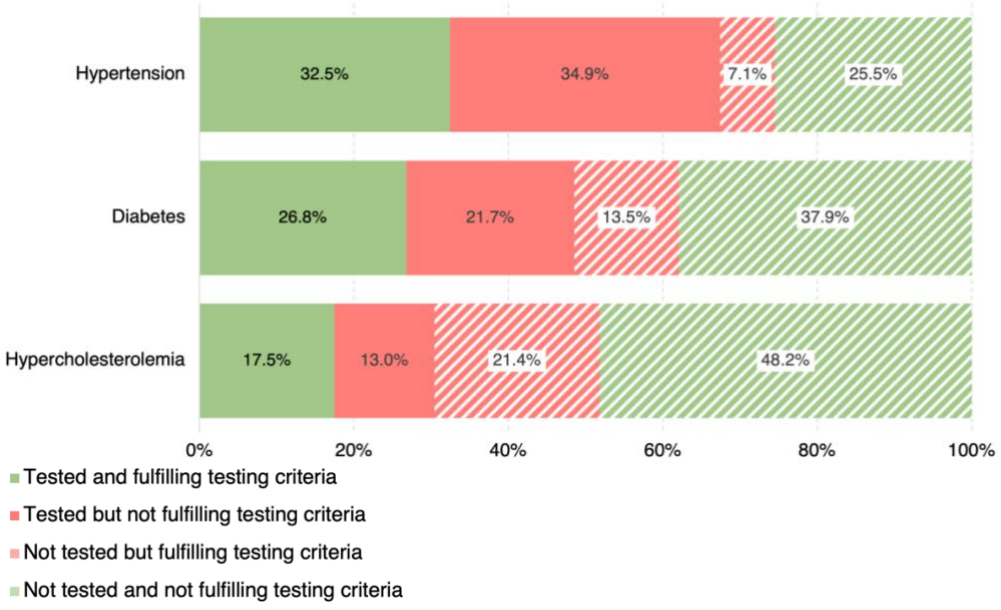
WHO World Region: Western Pacific




WHO World Region: Europe




WHO World Region: Eastern Mediterranean



*Appendix table A2.9: Diagnostic testing performance by World Bank Income Groups and WHO World Regions*

| CVD Risk factor | Dimension | Low-income countries (in %) | Lower-middle-income countries (in %) | Upper-middle-income countries (in %) |
|---|---|---|---|---|
| Hypertension | Fulfills diagnostic testing criteria | 13·9 (13·1-14·8) | 13·1 (12·9-13·4) | 43·5 (42·2-44·8) |
| | Tested | 38·3 (36·9-39·7) | 60·1 (59·6-60·5) | 83·3 (81·8-84·7) |
| | Tested out of all fulfilling criteria | 60·4 (58·1-62·7) | 77·5 (76·7-78·3) | 88·8 (87·6-89·8) |
| | Guidelines adhered | 64·8 (63·6-65·9) | 47·1 (46·7-47·5) | 50·6 (49·8-51·4) |
| Diabetes | Fulfills diagnostic testing criteria | 13·8 (13·0-14·7) | 16·9 (16·3-17·5) | 44·4 (43·7-45·1) |
| | Tested | 7·9 (7·3-8·5) | 24·2 (23·3-25·0) | 51·1 (50·0-52·1) |
| | Tested out of all fulfilling criteria | 21·1 (18·8-23·6) | 44·0 (41·9-46·1) | 62·3 (61·2-63·5) |
| | Guidelines adhered | 84·2 (83·5-84·9) | 73·9 (73·2-74·7) | 60·9 (60·3-61·5) |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 12·9 (11·7-14·2) | 19·1 (18·0-20·3) | 43·4 (42·0-44·9) |
| | Tested | 1·1 (0·8-1·5) | 11·4 (10·3-12·5) | 65·2 (61·9-68·3) |
| | Tested out of all fulfilling criteria | 3·4 (1·9-5·8) | 22·4 (20·3-24·6) | 73·6 (71·1-76·0) |
| | Guidelines adhered | 86·8 (85·5-88·1) | 78·2 (77·3-79·1) | 55·6 (55·0-56·1) |
| | | **Africa (in %)** | **Americas (in %)** | **South East Asia (in %)** |
| Hypertension | Fulfills diagnostic testing criteria | 22·6 (21·9-23·4) | 42·1 (41·3-42·8) | 12·1 (11·8-12·3) |
| | Tested | 45·1 (43·9-46·3) | 96·2 (96·0-96·5) | 60·5 (60·0-61·0) |
| | Tested out of all fulfilling criteria | 65·2 (63·5-66·9) | 98·1 (97·8-98·3) | 78·3 (77·4-79·1) |
| | Guidelines adhered | 62·4 (61·6-63·2) | 44·1 (43·4-44·9) | 46·4 (45·9-46·8) |
| Diabetes | Fulfills diagnostic testing criteria | 22·6 (21·9-23·3) | 45·4 (43·7-47·1) | 14·3 (13·3-15·3) |
| | Tested | 17·1 (16·4-17·9) | 48·5 (46·6-50·4) | 24·0 (22·7-25·4) |
| | Tested out of all fulfilling criteria | 30·9 (29·4-32·4) | 58·3 (55·9-60·6) | 48·3 (44·7-51·9) |
| | Guidelines adhered | 77·2 (76·5-77·9) | 59·0 (57·4-60·5) | 75·4 (74·2-76·5) |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 21·9 (21·0-22·9) | 42·1 (41·3-42·8) | 18·9 (17·2-20·8) |
| | Tested | 9·9 (9·2-10·7) | 82·2 (81·6-82·8) | 10·3 (8·7-12·3) |
| | Tested out of all fulfilling criteria | 15·6 (14·2-17·2) | 89·0 (88·4-89·6) | 22·0 (19·1-25·2) |
| | Guidelines adhered | 78·8 (77·9-79·7) | 50·4 (49·7-51·2) | 79·7 (78·1-81·3) |
| | | **Western Pacific (in %)** | **Europe (in %)** | **Eastern Mediterranean (in %)** |
| Hypertension | Fulfills diagnostic testing criteria | 12·0 (11·0-13·1) | 45·8 (38·0-53·8) | 39·6 (38·8-40·3) |
| | Tested | 65·0 (62·9-66·9) | 84·8 (72·4-92·2) | 67·4 (66·4-68·4) |
| | Tested out of all fulfilling criteria | 73·3 (69·0-77·2) | 89·7 (81·0-94·7) | 79·9 (78·8-81·0) |
| | Guidelines adhered | 40·5 (38·6-42·4) | 53·5 (50·3-56·6) | 57·9 (57·1-58·7) |
| Diabetes | Fulfills diagnostic testing criteria | 12·0 (11·0-13·1) | 39·8 (38·7-40·9) | 40·3 (39·6-41·1) |
| | Tested | 28·7 (27·0-30·6) | 42·6 (41·3-43·8) | 48·5 (47·5-49·5) |
| | Tested out of all fulfilling criteria | 42·8 (38·2-47·5) | 52·4 (50·7-54·1) | 61·8 (60·5-63·1) |
| | Guidelines adhered | 69·6 (67·9-71·4) | 62·4 (61·3-63·4) | 64·7 (64·0-65·5) |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 12·0 (10·8-13·3) | 45·8 (38·0-53·8) | 38·8 (38·1-39·6) |
| | Tested | 24·5 (22·5-26·6) | 51·6 (23·7-78·5) | 30·4 (29·8-31·1) |
| | Tested out of all fulfilling criteria | 41·4 (35·8-47·3) | 55·7 (28·6-79·7) | 40·7 (39·7-41·8) |
| | Guidelines adhered | 72·9 (70·8-74·8) | 61·2 (59·2-63·2) | 65·7 (65·0-66·3) |

Note: 95% confidence intervals in brackets.

# Appendix 2.12: Sensitivity analysis 1 (using equivalent weights)

*Appendix figure A2.9: WHO PEN diagnostic testing recommendations and testing status using equivalent weights*


*Appendix table A2.10: Diagnostic testing performance by CVD risk factor*

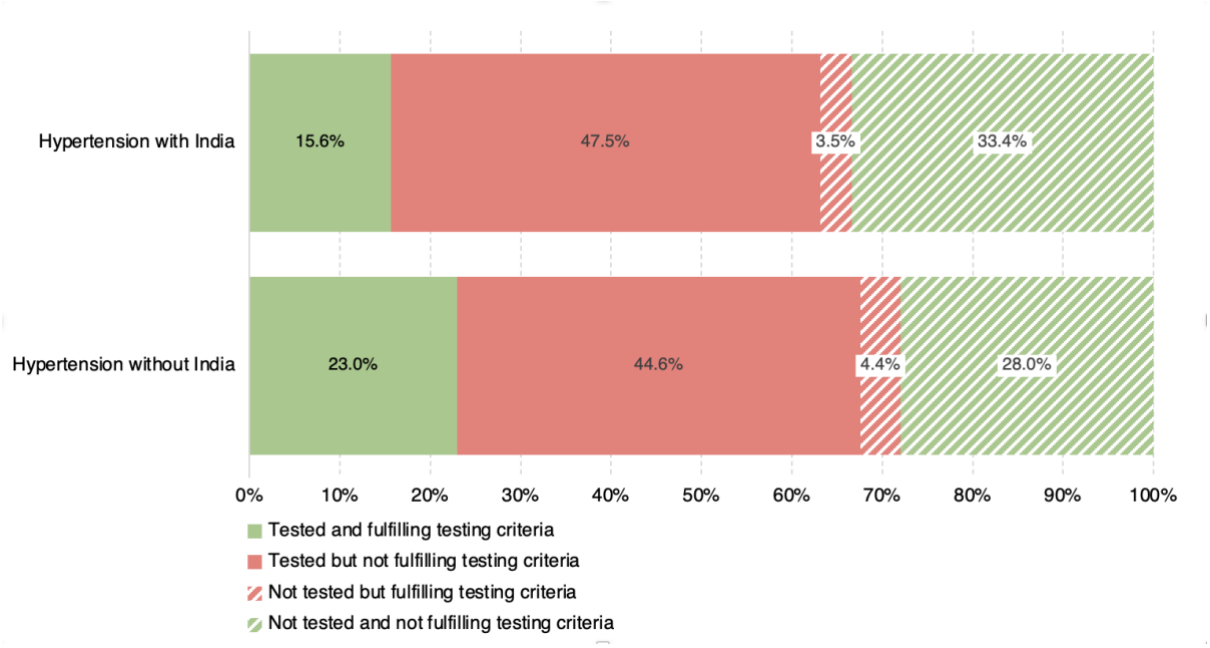
| CVD Risk factor | Dimension | Overall (in %) | | |
|---|---|---|---|---|
| Hypertension | Fulfills diagnostic testing criteria | 32·8 (31·3-34·4) | | |
| | Tested | 62·2 (60·7-63·8) | | |
| | Tested out of all fulfilling criteria | 74·3 (73·0-75·6) | | |
| | Guidelines adhered | 54·4 (53·9-54·9) | | |
| Diabetes | Fulfills diagnostic testing criteria | 33·0 (31·5-34·5) | | |
| | Tested | 30·6 (28·4-33·0) | | |
| | Tested out of all fulfilling criteria | 42·6 (40·6-44·7) | | |
| | Guidelines adhered | 69·3 (68·6-70·1) | | |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 36·6 (34·6-38·6) | | |
| | Tested | 23·0 (19·3-27·1) | | |
| | Tested out of all fulfilling criteria | 30·1 (26·6-34·0) | | |
| | Guidelines adhered | 66·5 (65·0-67·8) | | |
| **Sex** | | **Female (in %)** | **Male (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 37·8 (36·3-39·2) | 27·4 (25·7-29·3) | <0·0001* |
| | Tested | 68·5 (67·0-69·9) | 55·7 (53·9-57·4) | <0·0001* |
| | Tested out of all fulfilling criteria | 77·0 (75·7-78·3) | 69·9 (68·3-71·4) | <0·0001* |
| | Guidelines adhered | 52·3 (51·7-52·9) | 56·6 (56·0-57·3) | <0·0001* |
| Diabetes | Fulfills diagnostic testing criteria | 38·2 (36·7-39·8) | 27·3 (25·8-28·9) | <0·0001* |
| | Tested | 33·5 (31·1-35·9) | 27·8 (25·5-30·1) | 0·0008* |
| | Tested out of all fulfilling criteria | 43·4 (41·3-45·6) | 42·3 (40·1-44·4) | 0·4464 |

| | | | | | |
|---|---|---|---|---|---|
| | Guidelines adhered | 66·9 (66·3-67·5) | 71·9 (70·8-73·0) | <0·0001* | |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 41·4 (39·5-43·4) | 31·7 (29·4-34·0) | <0·0001* | |
| | Tested | 24·7 (21·0-28·8) | 21·2 (17·5-25·5) | 0·2182 | |
| | Tested out of all fulfilling criteria | 30·8 (27·2-34·6) | 29·6 (26·0-33·5) | 0·6591 | |
| | Guidelines adhered | 63·3 (61·9-64·7) | 69·6 (68·2-71·1) | <0·0001* | |
| **Wealth** | | **Poorest quintile (in %)** | **Middle quintile (in %)** | **Richest quintile (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 28·3 (26·6-30·0) | 31·5 (29·6-33·6) | 37·2 (35·0-39·6) | 0·0001* |
| | Tested | 55·2 (53·0-57·4) | 59·7 (57·6-61·8) | 67·4 (65·4-69·4) | <0·0001* |
| | Tested out of all fulfilling criteria | 66·3 (63·9-68·6) | 72·6 (70·5-74·6) | 78·7 (77·0-80·3) | <0·0001* |
| | Guidelines adhered | 56·5 (55·3-57·8) | 55·2 (54·2-56·1) | 54·2 (52·7-55·6) | 0·0862 |
| Diabetes | Fulfills diagnostic testing criteria | 29·3 (27·1-31·6) | 31·7 (29·8-33·6) | 37·7 (35·0-40·4) | 0·0001* |
| | Tested | 22·8 (19·6-26·3) | 27·8 (24·7-31·0) | 37·1 (33·8-40·6) | <0·0001* |
| | Tested out of all fulfilling criteria | 32·5 (29·2-35·9) | 39·1 (36·1-42·2) | 49·4 (46·3-52·4) | <0·0001* |
| | Guidelines adhered | 72·4 (71·3-73·5) | 70·1 (68·8-71·4) | 66·4 (65·3-67·6) | <0·0001* |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 34·4 (31·7-37·2) | 36·7 (34·2-39·4) | 40·7 (37·1-44·3) | 0·0529 |
| | Tested | 18·4 (13·5-24·7) | 22·2 (17·4-27·7) | 28·0 (23·1-33·5) | 0·0198* |
| | Tested out of all fulfilling criteria | 24·3 (19·3-30·1) | 28·7 (24·0-33·8) | 34·9 (30·3-39·8) | 0·0219* |
| | Guidelines adhered | 66·8 (64·3-69·2) | 65·9 (64·0-67·6) | 62·7 (60·5-64·9) | 0·0120* |
| **Education** | | **Less than primary school (in %)** | **Less than secondary school (in %)** | **Secondary school or more (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 33·4 (30·9-36·0) | 32·3 (30·4-34·4) | 33·1 (31·0-35·2) | 0·8457 |
| | Tested | 60·0 (56·9-63·1) | 60·8 (59·1-62·4) | 62·1 (60·3-63·9) | 0·2597 |
| | Tested out of all fulfilling criteria | 67·9 (64·1-71·5) | 73·6 (71·2-76·0) | 74·9 (73·0-76·7) | 0·0010* |
| | Guidelines adhered | 55·2 (52·9-57·6) | 55·4 (53·7-57·2) | 55·1 (53·9-56·3) | 0·9260 |
| Diabetes | Fulfills diagnostic testing criteria | 33·2 (30·7-35·8) | 32·3 (30·2-34·4) | 33·1 (31·2-35·1) | 0·9886 |
| | Tested | 24·2 (21·1-27·6) | 25·7 (23·9-27·5) | 28·4 (25·7-31·2) | 0·0539 |
| | Tested out of all fulfilling criteria | 35·5 (31·4-39·8) | 36·2 (33·8-38·7) | 39·9 (37·2-42·7) | 0·0849 |
| | Guidelines adhered | 72·1 (69·3-74·8) | 71·3 (69·6-72·9) | 69·6 (68·4-70·7) | 0·0938 |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 37·8 (34·5-41·2) | 35·3 (32·9-37·8) | 37·6 (34·8-40·5) | 0·9534 |
| | Tested | 19·6 (14·6-25·8) | 17·0 (15·7-18·4) | 20·7 (16·6-25·5) | 0·7601 |
| | Tested out of all fulfilling criteria | 25·4 (19·6-32·2) | 22·8 (20·6-25·2) | 26·8 (22·7-31·4) | 0·7181 |
| | Guidelines adhered | 68·4 (64·4-72·1) | 68·3 (65·7-70·8) | 66·4 (64·4-68·3) | 0·3606 |

Note: 95% confidence intervals in brackets. The p-values were calculated using an immediate form of a two-sample t-test with unequal variances. We compared male vs. female sex, less than primary vs. secondary or more education, and richest vs. remaining four wealth quintiles. Statistical significance levels based on Benjamini-Hochberg adjusted p-values corresponding to a 5% significance level represented with *.

# Appendix 2.13: Sensitivity analysis 2 (hypertension analysis excluding India)

*Appendix figure A2.10: WHO PEN diagnostic testing recommendations and testing status excluding India*



*Appendix table A2.11: Diagnostic testing performance by CVD risk factor when excluding India*

| CVD Risk factor | Dimension | Overall (in %) | | | |
|---|---|---|---|---|---|
| Hypertension | Fulfills diagnostic testing criteria | 27·4 (26·3-28·5) | | | |
| | Tested | 67·6 (66·3-68·8) | | | |
| | Tested out of all fulfilling criteria | 80·2 (79·0-81·2) | | | |
| | Guidelines adhered | 51·0 (50·4-51·5) | | | |
| **Sex** | | **Female (in %)** | **Male (in %)** | **p-value from two-sided t-test** | |
| Hypertension | Fulfills diagnostic testing criteria | 32·1 (31·2-33·0) | 22·3 (21·0-23·7) | <0·0001* | |
| | Tested | 74·6 (73·5-75·7) | 60·4 (58·8-62·0) | <0·0001* | |
| | Tested out of all fulfilling criteria | 82·9 (81·7-84·0) | 75·1 (73·5-76·7) | <0·0001* | |
| | Guidelines adhered | 48·1 (47·4-48·7) | 53·9 (53·1-54·7) | <0·0001* | |
| **Wealth** | | **Poorest quintile (in %)** | **Middle quintile (in %)** | **Richest quintile (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 25·3 (24·3-26·3) | 31·3 (29·9-32·8) | 37·4 (35·7-39·1) | <0·0001* |
| | Tested | 60·4 (58·3-62·5) | 66·7 (64·8-68·5) | 73·5 (70·8-76·0) | <0·0001* |
| | Tested out of all fulfilling criteria | 72·4 (70·0-74·7) | 79·2 (77·3-81·0) | 83·8 (81·8-85·6) | <0·0001* |
| | Guidelines adhered | 54·6 (53·4-55·8) | 54·3 (53·3-55·2) | 53·4 (52·0-54·9) | 0·0542 |
| **Education** | | **Less than primary school (in %)** | **Less than secondary school (in %)** | **Secondary school or more (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 31·6 (30·2-33·1) | 32·7 (31·6-33·7) | 28·9 (27·2-30·5) | 0·0144* |
| | Tested | 65·4 (63·5-67·3) | 67·5 (66·2-68·7) | 67·1 (65·4-68·8) | 0·1965 |
| | Tested out of all fulfilling criteria | 73·5 (70·5-76·3) | 79·8 (77·6-81·8) | 79·6 (77·2-81·8) | 0·0014* |
| | Guidelines adhered | 54·3 (52·9-55·8) | 55·9 (54·7-57·0) | 52·2 (50·8-53·6) | 0·0351* |

Note: 95% confidence intervals in brackets. The p-values were calculated using an immediate form of a two-sample t-test with unequal variances. We compared male vs. female sex, less than primary vs. secondary or more education, and richest vs. remaining four wealth quintiles. Statistical significance levels based on Benjamini-Hochberg adjusted p-values corresponding to a 5% significance level represented with *.

# Appendix 2.14: Sensitivity analysis 3 (testing performance below and above 40 years old)

*Appendix table A2.12: Diagnostic testing performance by CVD risk factor for individuals aged 18-39 years (young) vs. individuals aged 40+ years (old)*

| CVD Risk factor | Dimension | Young (in %) | Old (in %) | |
|---|---|---|---|---|
| Hypertension | Fulfills diagnostic testing criteria | 6·8 (6·6-7·0) | 38·0 (37·2-38·8) | |
| | Tested | 59·2 (58·4-60·0) | 69·9 (69·2-70·6) | |
| | Tested out of all fulfilling criteria | 75·7 (74·1-77·2) | 79·4 (78·6-80·1) | |
| | Guidelines adhered | 44·5 (43·9-45·2) | 54·2 (53·6-54·8) | |
| Diabetes | Fulfills diagnostic testing criteria | 9·1 (8·7-9·4) | 41·8 (41·1-42·6) | |
| | Tested | 22·1 (21·5-22·8) | 36·7 (35·9-37·6) | |
| | Tested out of all fulfilling criteria | 37·2 (33·3-41·3) | 47·4 (46·0-48·8) | |
| | Guidelines adhered | 76·2 (75·5-76·8) | 65·8 (65·1-66·5) | |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 9·7 (9·3-10·1) | 46.4 (44·6-48·1) | |
| | Tested | 24·5 (21·6-27·8) | 35.3 (32·4-38·4) | |
| | Tested out of all fulfilling criteria | 34·1 (30·0-38·4) | 41.2 (38·6-43·9) | |
| | Guidelines adhered | 74·1 (71·6-76·5) | 64.5 (63·7-65·4) | |
| **Sex** | | **Young females (in %)** | **Young males (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 8·7 (8·4-8·9) | 4·9 (4·6-5·2) | <0·0001* |
| | Tested | 67·8 (67·1-68·5) | 50·5 (49·5-51·5) | <0·0001* |
| | Tested out of all fulfilling criteria | 81·2 (79·8-82·6) | 66·6 (63·4-69·7) | <0·0001* |
| | Guidelines adhered | 37·4 (36·8-38·0) | 51·7 (50·9-52·5) | <0·0001* |
| Diabetes | Fulfills diagnostic testing criteria | 12·2 (11·7-12·7) | 5·6 (5·2-6·1) | <0·0001* |
| | Tested | 25·2 (24·4-26·1) | 19·0 (18·2-19·9) | <0·0001* |
| | Tested out of all fulfilling criteria | 39·9 (33·3-46·9) | 35·8 (29·9-42·2) | 0·3857 |
| | Guidelines adhered | 72·5 (71·7-73·3) | 80·0 (79·1-80·9) | <0·0001* |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 12·5 (12·1-13·0) | 6·9 (6·3-7·6) | <0·0001* |
| | Tested | 26·4 (23·4-29·8) | 22·6 (19·7-25·8) | 0·0904 |
| | Tested out of all fulfilling criteria | 35·7 (29·7-42·2) | 33·0 (28·3-38·2) | 0·5133 |
| | Guidelines adhered | 71·1 (68·5-73·5) | 77·1 (74·5-79·5) | 0·0008* |
| **Sex** | | **Old females (in %)** | **Old males (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 42·9 (42·2-43·6) | 33·5 (32·5-34·6) | <0·0001* |
| | Tested | 74·8 (74·2-75·3) | 65·7 (64·9-66·6) | <0·0001* |
| | Tested out of all fulfilling criteria | 82·5 (81·8-83·1) | 75·6 (74·5-76·7) | <0·0001* |
| | Guidelines adhered | 54·5 (53·9-55·0) | 53·8 (53·0-54·6) | 0.1667 |
| Diabetes | Fulfills diagnostic testing criteria | 49·4 (48·4-50·4) | 33·9 (32·8-35·0) | <0·0001* |
| | Tested | 39·5 (38·4-40·5) | 33·9 (32·7-35·0) | <0·0001* |
| | Tested out of all fulfilling criteria | 48·3 (46·5-50·0) | 46·4 (44·5-48·3) | 0·1579 |
| | Guidelines adhered | 63·7 (62·8-64·6) | 68·1 (67·1-69·1) | <0·0001* |

| | | | | | |
|---|---|---|---|---|---|
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 53·1 (51·7-54·4) | 39·5 (37·0-42·0) | <0·0001* | |
| | Tested | 37·0 (34·1-40·1) | 33·4 (30·4-36·5) | 0·0935 | |
| | Tested out of all fulfilling criteria | 41·3 (38·5-44·1) | 41·6 (38·9-44·3) | 0·8998 | |
| | Guidelines adhered | 60·8 (59·8-61·7) | 68·3 (67·1-69·5) | <0·0001* | |
| **Wealth** | | **Young in poorest quintile (in %)** | **Young in middle quintile (in %)** | **Young in richest quintile (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 4·0 (3·7-4·3) | 6·5 (6·1-6·8) | 9·6 (9·1-10·1) | <0·0001* |
| | Tested | 45.8 (44.6-47) | 57·8 (56·8-58·9) | 68·3 (67·1-69·5) | <0·0001* |
| | Tested out of all fulfilling criteria | 59 (54.1-63.8) | 70·9 (68·1-73·7) | 79·0 (76·8-81·2) | <0·0001* |
| | Guidelines adhered | 55.9 (54.8-56.9) | 45·3 (44·4-46·2) | 37·5 (36·4-38·5) | <0·0001* |
| Diabetes | Fulfills diagnostic testing criteria | 8·4 (7·7-9·1) | 10·7 (9·9-11·6) | 12·4 (11·3-13·6) | 0·0014* |
| | Tested | 14·2 (13·2-15·2) | 20·0 (18·9-21·1) | 28·8 (26·9-30·7) | <0·0001* |
| | Tested out of all fulfilling criteria | 24·4 (19·7-29·8) | 30·6 (26·3-35·3) | 41·0 (36·1-46·2) | <0·0001* |
| | Guidelines adhered | 82·7 (81·7-83·7) | 76·9 (75·7-78·0) | 69·8 (67·7-71·8) | <0·0001* |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 8·7 (8·0-9·5) | 11·6 (10·6-12·6) | 12·7 (11·8-13·5) | 0·0011* |
| | Tested | 22·3 (18·3-26·9) | 29·7 (24·8-35·2) | 37·6 (32·6-42·8) | 0·0001* |
| | Tested out of all fulfilling criteria | 29·4 (24·2-35·1) | 37·3 (32·3-42·6) | 43·1 (37·4-48·9) | 0·0471 |
| | Guidelines adhered | 75·6 (71·8-79·0) | 69·7 (65·2-73·8) | 62·7 (58·9-66·3) | 0·0020* |
| **Wealth** | | **Old in poorest quintile (in %)** | **Old in middle quintile (in %)** | **Old in richest quintile (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 21·6 (20·4-22·8) | 36·5 (35·4-37·6) | 55·2 (53·8-56·5) | <0·0001* |
| | Tested | 53·0 (51·8-54·2) | 68·2 (67·2-69·3) | 82·0 (80·5-83·3) | <0·0001* |
| | Tested out of all fulfilling criteria | 62·7 (60·3-65·0) | 75·7 (74·1-77·2) | 85·4 (84·3-86·5) | <0·0001* |
| | Guidelines adhered | 57·5 (56·6-58·3) | 53·0 (52·1-54·0) | 58·0 (56·8-59·1) | <0·0001* |
| Diabetes | Fulfills diagnostic testing criteria | 37·3 (36·0-38·6) | 47·3 (45·7-48·9) | 59·1 (55·9-62·2) | <0·0001* |
| | Tested | 26·2 (25·1-27·4) | 35·3 (33·8-36·9) | 47·6 (44·7-50·5) | <0·0001* |
| | Tested out of all fulfilling criteria | 33·1 (30·7-35·7) | 43·7 (40·4-47·1) | 54·6 (52·0-57·1) | <0·0001* |
| | Guidelines adhered | 67·9 (66·7-69·1) | 64·7 (63·2-66·1) | 62·4 (59·8-65·0) | <0·0001* |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 40·3 (36·7-44·0) | 52·1 (49·0-55·2) | 63·3 (60·9-65·7) | <0·0001* |
| | Tested | 32·7 (28·1-37·6) | 41·7 (36·4-47·2) | 53·0 (48·8-57·1) | <0·0001* |
| | Tested out of all fulfilling criteria | 37·5 (33·6-41·6) | 46·0 (41·2-50·9) | 55·6 (52·0-59·2) | <0·0001* |
| | Guidelines adhered | 64·9 (62·3-67·4) | 62·4 (61·1-63·8) | 57·6 (55·8-59·4) | <0·0001* |
| **Education** | | **Young with less than primary school (in %)** | **Young with less than secondary school (in %)** | **Young with secondary school or more (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 6·4 (5·6-7·5) | 6.3 (5.6-7·0) | 7·0 (6·4-7·6) | 0·3429 |
| | Tested | 54·2 (52·6-55·7) | 56·6 (55·3-58·0) | 58·1 (56·8-59·4) | 0·0001* |
| | Tested out of all fulfilling criteria | 68·8 (65·3-72·1) | 75·2 (68·9-80·6) | 74·4 (69·0-79·1) | 0·0730 |
| | Guidelines adhered | 49·0 (47·6-50·4) | 47·3 (46·0-48·7) | 45·2 (44·0-46·4) | <0·0001* |
| Diabetes | Fulfills diagnostic testing criteria | 8·8 (7·6-10·2) | 9·5 (8·2-11·0) | 10·0 (8.8-11·3) | 0·2113 |
| | Tested | 16·8 (15·0-18·8) | 17·7 (16·1-19·4) | 19·3 (17·7-20·9) | 0·0558 |

| | | | | | |
|---|---|---|---|---|---|
| | Tested out of all fulfilling criteria | 31·8 (22·6-42·6) | 35·3 (23·3-49·6) | 28·8 (24·6-33·3) | 0·5921 |
| | Guidelines adhered | 81·6 (79·7-83·4) | 81·4 (79·8-82·9) | 77·7 (75·9-79·4) | 0·0024* |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 10·6 (8·2-13·7) | 10·1 (8·7-11·7) | 10·6 (9·7-11·5) | 0·9733 |
| | Tested | 18·8 (15·4-22·8) | 18·8 (16·4-21·4) | 21·3 (17·9-25·2) | 0·3405 |
| | Tested out of all fulfilling criteria | 24·2 (16·6-33·8) | 28·2 (18·9-39·8) | 26·7 (22·9-30·8) | 0·6041 |
| | Guidelines adhered | 79·6 (76·4-82·5) | 79·0 (76·4-81·3) | 76·3 (73·3-79·1) | 0·1170 |
| **Education** | | **Old with less than primary school (in %)** | **Old with less than secondary school (in %)** | **Old with secondary school or more (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 29·2 (28·3-30·1) | 34·0 (32·9-35·1) | 39·4 (38·0-40·8) | <0·0001* |
| | Tested | 62·8 (61·8-63·8) | 67·6 (66·3-68·8) | 69·6 (68·3-70·9) | <0·0001* |
| | Tested out of all fulfilling criteria | 72·9 (71·2-74·5) | 75·6 (73·2-77·9) | 78·4 (76·4-80·3) | <0·0001* |
| | Guidelines adhered | 52·9 (52·1-53·8) | 52·3 (51·1-53·5) | 54·7 (53·4-56·0) | 0·0212* |
| Diabetes | Fulfills diagnostic testing criteria | 36·1 (34·6-37·8) | 40·5 (38·7-42·3) | 43·8 (42·0-45·5) | <0·0001* |
| | Tested | 28·8 (27·2-30·5) | 34·0 (32·1-36·0) | 37·3 (35·6-39·1) | <0·0001* |
| | Tested out of all fulfilling criteria | 38·5 (35·1-42·1) | 41·8 (38·7-45·0) | 48·1 (45·1-51·1) | <0·0001* |
| | Guidelines adhered | 69·1 (67·4-70·7) | 65·1 (63·3-66·9) | 65·0 (63·3-66·8) | 0·0008* |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 37·5 (34·3-40·9) | 43·6 (41·5-45·7) | 47·9 (45·4-50·5) | <0·0001* |
| | Tested | 27·9 (24·1-32·1) | 30·0 (27·4-32·8) | 32·9 (29·2-36·9) | 0·0751 |
| | Tested out of all fulfilling criteria | 31·4 (27·9-35·2) | 34·8 (32·1-37·6) | 39·3 (35·5-43·2) | 0·0042* |
| | Guidelines adhered | 67·2 (64·9-69·5) | 65·4 (63·9-66·9) | 64·3 (62·7-65·9) | 0·0399* |

Note: 95% confidence intervals in brackets. The p-values were calculated using an immediate form of a two-sample t-test with unequal variances. We compared male vs. female sex, less than primary vs. secondary or more education, and richest vs. remaining four wealth quintiles. Statistical significance levels based on Benjamini-Hochberg adjusted p-values corresponding to a 5% significance level represented with *.

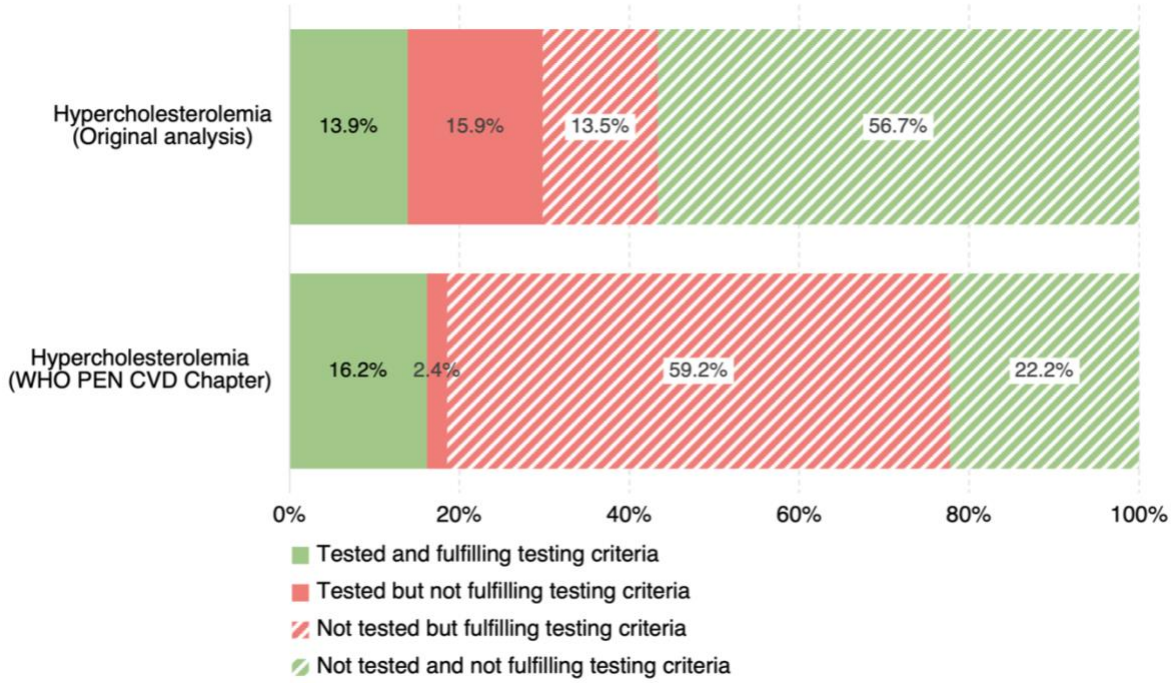
## Appendix 2.15: Sensitivity analysis 4 (hypercholesterolemia analysis using CVD chapter of WHO PEN guidelines)

Methodological note:

The CVD chapter of the WHO PEN guidelines[3] recommends testing individuals fulfilling one of the following criteria:

- aged over 40 years
- has a history of tobacco use
- is overweight
- has hypertension
- has diabetes mellitus
- has a history of premature CVD in first degree relatives
- has a history of diabetes mellitus or kidney disease in first-degree relatives

*Appendix figure A2.11: WHO PEN diagnostic testing recommendations and testing status using the CVD chapter*



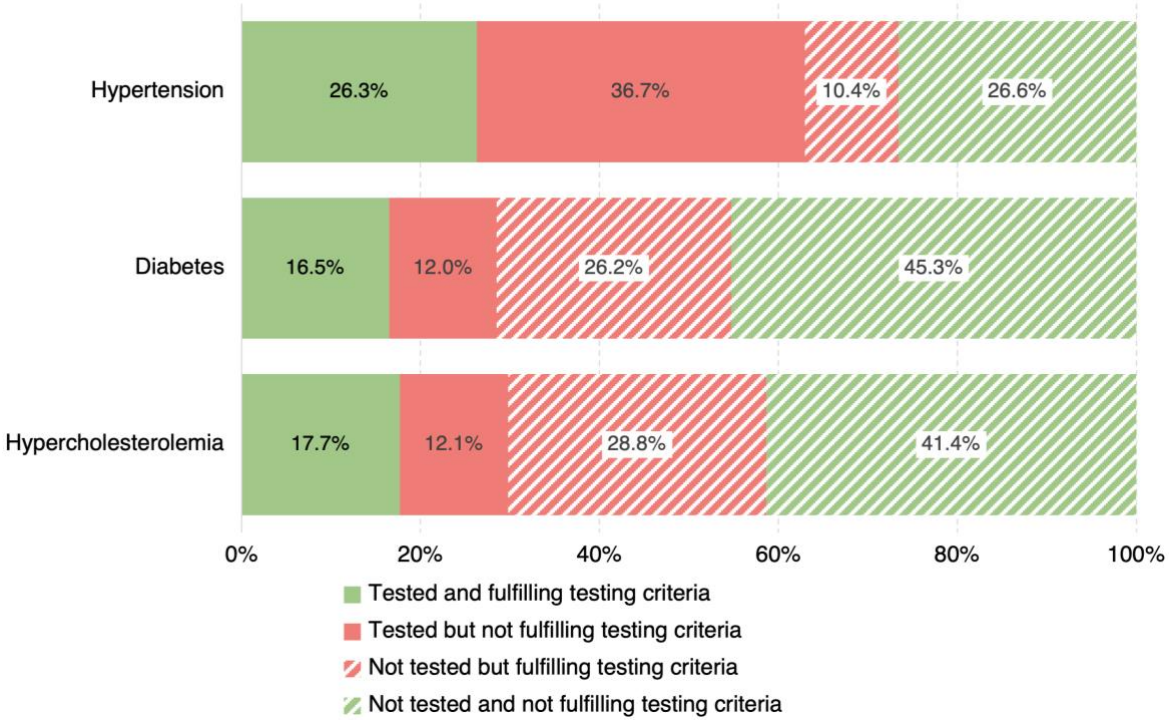
*Appendix table A2.13: Diagnostic testing performance by CVD risk factor when using WHO PEN guidelines' chapter 2.1*

| CVD Risk factor | Dimension | Overall (in %) | | | |
|---|---|---|---|---|---|
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 75·4 (74·3-76·4) | | | |
| | Tested | 18·5 (15·5-22·0) | | | |
| | Tested out of all fulfilling criteria | 20·3 (17·2-23·7) | | | |
| | Guidelines adhered | 38·4 (36·1-40·8) | | | |
| **Sex** | | **Female (in %)** | **Male (in %)** | **p-value from two-sided t-test** | |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 75·3 (74·2-76·4) | 81·3 (80·2-82·4) | <0·0001* | |
| | Tested | 24·4 (20·7-28·5) | 20·4 (16·7-24·8) | 0·2479 | |
| | Tested out of all fulfilling criteria | 26·6 (22·9-30·6) | 21·8 (18·1-26·0) | 0·0798 | |
| | Guidelines adhered | 43·3 (40·9-45·6) | 35·7 (32·8-38·7) | <0·0001* | |
| **Wealth** | | **Poorest quintile (in %)** | **Middle quintile (in %)** | **Richest quintile (in %)** | **p-value from two-sided t-test** |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 78·6 (76·8-80·2) | 79·0 (76·9-80·9) | 79·4 (77·2-81·4) | 0·6549 |
| | Tested | 16·8 (11·7-23·5) | 20·1 (15·1-26·1) | 25·6 (20·4-31·6) | 0·0093* |
| | Tested out of all fulfilling criteria | 18·1 (13·0-24·7) | 21·0 (16·0-27·0) | 27·7 (22·5-33·5) | 0·0042* |
| | Guidelines adhered | 34·6 (31·3-38·1) | 36·2 (32·3-40·2) | 40·9 (37·2-44·8) | 0·0119* |
| **Education** | | **Less than primary school (in %)** | **Less than secondary school (in %)** | **Secondary school or more (in %)** | **p-value from two-sided t-test** |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 87·1 (84·5-89·4) | 80·8 (79·0-82·5) | 80·0 (77·9-82·0) | <0·0001* |
| | Tested | 18·5 (14·1-24·0) | 16·4 (14·9-17·9) | 20·4 (16·3-25·3) | 0·6456 |
| | Tested out of all fulfilling criteria | 18·5 (14·1-23·9) | 17·5 (15·9-19·2) | 21·5 (17·4-26·2) | 0·5435 |
| | Guidelines adhered | 29·2 (24·6-34·2) | 33·8 (31·6-36·1) | 37·0 (33·6-40·5) | 0·0014* |

Note: 95% confidence intervals in brackets. The p-values were calculated using an immediate form of a two-sample t-test with unequal variances. We compared male vs. female sex, less than primary vs. secondary or more education, and richest vs. remaining four wealth quintiles. Statistical significance levels based on Benjamini-Hochberg adjusted p-values corresponding to a 5% significance level represented with *.

# Appendix 2.16: Sensitivity analysis 5 (using AHA/ACC guidelines)

*Appendix figure A2.12: AHA/ACC diagnostic testing recommendations and testing status*




*Appendix table A2.14: Diagnostic testing performance by CVD risk factor*

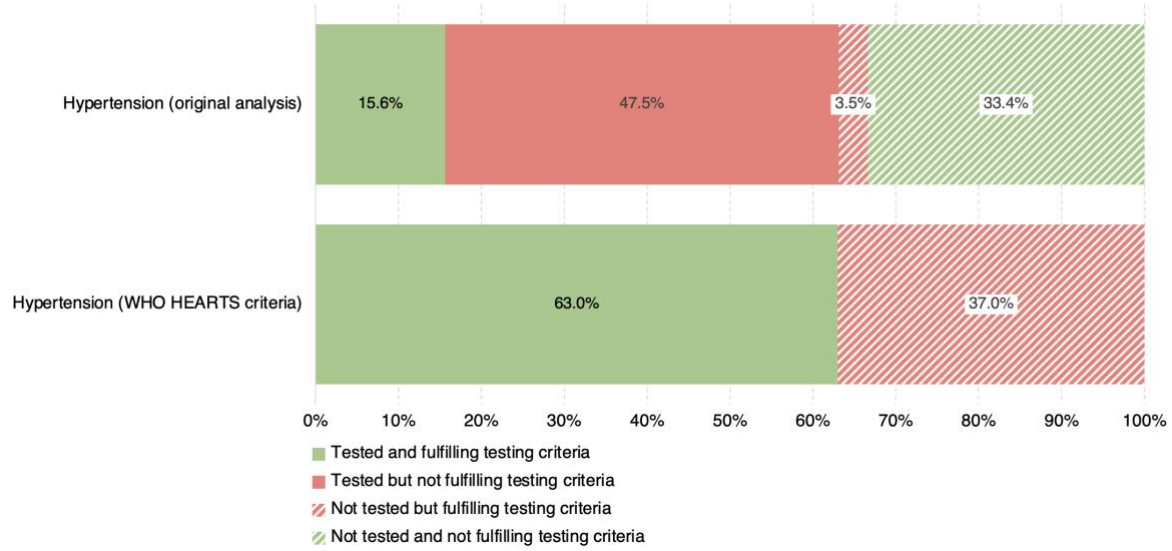
| CVD Risk factor | Dimension | Overall (in %) | | |
|---|---|---|---|---|
| Hypertension | Fulfills diagnostic testing criteria | 36·7 (36·2-37·2) | | |
| | Tested | 63·0 (62·3-63·7) | | |
| | Tested out of all fulfilling criteria | 69·8 (69·1-70·5) | | |
| | Guidelines adhered | 52·9 (52·6-53·2) | | |
| Diabetes | Fulfills diagnostic testing criteria | 42·7 (42·1-43·3) | | |
| | Tested | 28·5 (27·9-29·1) | | |
| | Tested out of all fulfilling criteria | 36·8 (36·0-37·6) | | |
| | Guidelines adhered | 61·8 (61·2-62·3) | | |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 46·5 (45·3-47·8) | | |
| | Tested | 29·8 (26·9-32·9) | | |
| | Tested out of all fulfilling criteria | 35·3 (32·4-38·4) | | |
| | Guidelines adhered | 59·1 (58·1-60·0) | | |
| **Sex** | | **Female (in %)** | **Male (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 35·0 (34·5-35·6) | 38·2 (37·7-38·8) | <0·0001* |
| | Tested | 70·2 (69·6-70·8) | 56·1 (55·2-57·0) | <0·0001* |
| | Tested out of all fulfilling criteria | 74·7 (74·1-75·2) | 65·6 (64·7-66·5) | <0·0001* |

| | | | | | |
|---|---|---|---|---|---|
| | Guidelines adhered | 48·5 (48·1-48·9) | 57·0 (56·5-57·4) | <0·0001* | |
| Diabetes | Fulfills diagnostic testing criteria | 42·9 (42·2-43·7) | 42·6 (41·8-43·3) | 0·4970 | |
| | Tested | 31·6 (30·9-32·4) | 25·4 (24·7-26·2) | <0·0001* | |
| | Tested out of all fulfilling criteria | 39·6 (38·5-40·6) | 34·0 (32·9-35·1) | <0·0001* | |
| | Guidelines adhered | 61·3 (60·7-62·0) | 62·2 (61·4-62·9) | 0·0955 | |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 47·3 (46·0-48·6) | 45·8 (44·5-47·1) | 0·1065 | |
| | Tested | 31·8 (28·9-34·9) | 27·7 (24·9-30·8) | 0·0609 | |
| | Tested out of all fulfilling criteria | 37·1 (34·2-40·1) | 33·4 (30·4-36·5) | 0·0883 | |
| | Guidelines adhered | 59·2 (58·2-60·2) | 58·9 (57·8-60·0) | 0·7088 | |
| **Wealth** | | **Poorest quintile (in %)** | **Middle quintile (in %)** | **Richest quintile (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 36·2 (35·8-36·7) | 34·9 (34·1-35·6) | 38·2 (37·4-39·0) | <0·0001* |
| | Tested | 48·6 (47·5-49·7) | 61·3 (60·4-62·2) | 72·9 (71·8-74·1) | <0·0001* |
| | Tested out of all fulfilling criteria | 52·9 (51·7-54·1) | 68·1 (67·0-69·1) | 81·8 (80·4-83·1) | <0·0001* |
| | Guidelines adhered | 56·2 (55·6-56·7) | 52·8 (52·3-53·3) | 51·7 (51·1-52·3) | <0·0001* |
| Diabetes | Fulfills diagnostic testing criteria | 45·4 (44·4-46·4) | 42·6 (41·5-43·8) | 43·8 (42·2-45·4) | 0·7207 |
| | Tested | 19·7 (18·9-20·6) | 26·4 (25·5-27·4) | 36·2 (34·5-37·9) | <0·0001* |
| | Tested out of all fulfilling criteria | 26·1 (24·9-27·2) | 35·2 (33·7-36·7) | 47·0 (44·2-49·8) | <0·0001* |
| | Guidelines adhered | 60·4 (59·4-61·3) | 62·6 (61·5-63·6) | 62·8 (61·1-64·5) | 0·4748 |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 50·2 (48·5-51·8) | 48·9 (47·2-50·6) | 48·8 (47·3-50·4) | 0·8424 |
| | Tested | 27·7 (23·4-32·3) | 35·6 (30·6-41·0) | 44·9 (40·4-49·5) | 0·0009* |
| | Tested out of all fulfilling criteria | 32·6 (28·0-37·5) | 41·7 (36·4-47·2) | 53·0 (48·8-57·2) | <0·0001* |
| | Guidelines adhered | 55·3 (53·3-57·2) | 57·2 (55·7-58·7) | 61·5 (60·5-62·5) | <0·0001* |
| **Education** | | **Less than primary school (in %)** | **Less than secondary school (in %)** | **Secondary school or more (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 56·7 (55·9-57·4) | 48·6 (47·8-49·4) | 38·8 (37·8-39·7) | <0·0001* |
| | Tested | 59·4 (58·4-60·4) | 61·9 (61·0-62·8) | 62·6 (61·6-63·6) | <0·0001* |
| | Tested out of all fulfilling criteria | 62·6 (61·6-63·7) | 67·5 (66·3-68·7) | 69·5 (68·2-70·7) | <0·0001* |
| | Guidelines adhered | 58·3 (57·5-59·0) | 58·4 (57·7-59·2) | 53·7 (52·9-54·5) | <0·0001* |
| Diabetes | Fulfills diagnostic testing criteria | 65·3 (63·7-66·9) | 51·4 (50·0-52·8) | 42·6 (41·2-44·0) | <0·0001* |
| | Tested | 25·1 (23·8-26·4) | 26·3 (25·1-27·5) | 26·8 (25·7-28·0) | 0·0429 |
| | Tested out of all fulfilling criteria | 28·9 (27·3-30·5) | 33·8 (32·0-35·7) | 37·4 (35·7-39·1) | <0·0001* |
| | Guidelines adhered | 51·1 (49·6-52·5) | 60·0 (58·5-61·4) | 64·6 (63·2-65·9) | <0·0001* |
| Hyper-cholesterolemia | Fulfills diagnostic testing criteria | 74·2 (71·0-77·2) | 64·8 (61·6-67·8) | 47·5 (45·2-49·9) | <0·0001* |
| | Tested | 25·7 (22·1-29·6) | 27·1 (24·9-29·4) | 27·2 (23·7-30·9) | 0·5658 |
| | Tested out of all fulfilling criteria | 27·7 (23·9-32·0) | 29·9 (27·4-32·6) | 32·8 (29·1-36·8) | 0·0731 |
| | Guidelines adhered | 44·6 (39·8-49·6) | 56·1 (51·8-60·4) | 60·6 (58·1-63·0) | <0·0001* |

Note: 95% confidence intervals in brackets. The p-values were calculated using an immediate form of a two-sample t-test with unequal variances. We compared male vs. female sex, less than primary vs. secondary or more education, and richest vs. remaining four wealth quintiles. Statistical significance levels based on Benjamini-Hochberg adjusted p-values corresponding to a 5% significance level represented with *.

# Appendix 2.17: Sensitivity analysis 6 (hypertension analysis using WHO HEARTS guidelines)

*Appendix figure A2.13: WHO HEARTS vs. WHO PEN diagnostic testing recommendations and testing status*



*Appendix table A2.15: Diagnostic testing performance by CVD risk factor when using WHO HEARTS*

| CVD Risk factor | Dimension | Overall (in %) | | | |
|---|---|---|---|---|---|
| Hypertension | Fulfills diagnostic testing criteria | 100·0 | | | |
| | Tested | 63·0 (62·3-63·7) | | | |
| | Tested out of all fulfilling criteria | 63·0 (62·3-63·7) | | | |
| | Guidelines adhered | 63·0 (62·3-63·7) | | | |
| **Sex** | | **Female (in %)** | **Male (in %)** | **p-value from two-sided t-test** | |
| Hypertension | Fulfills diagnostic testing criteria | 100·0 | 100·0 | | |
| | Tested | 70·3 (69·7-70·8) | 56·2 (55·3-57·0) | <0·0001* | |
| | Tested out of all fulfilling criteria | 70·3 (69·7-70·8) | 56·2 (55·3-57·0) | <0·0001* | |
| | Guidelines adhered | 70·3 (69·7-70·8) | 56·2 (55·3-57·0) | <0·0001* | |
| **Wealth** | | **Poorest quintile (in %)** | **Middle quintile (in %)** | **Richest quintile (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 100·0 | 100·0 | 100·0 | |
| | Tested | 48·7 (47·6-49·8) | 61·3 (60·4-62·3) | 73·0 (71·8-74·1) | <0·0001* |
| | Tested out of all fulfilling criteria | 48·7 (47·6-49·8) | 61·3 (60·4-62·3) | 73·0 (71·8-74·1) | <0·0001* |
| | Guidelines adhered | 48·7 (47·6-49·8) | 61·3 (60·4-62·3) | 73·0 (71·8-74·1) | <0·0001* |
| **Education** | | **Less than primary school (in %)** | **Less than secondary school (in %)** | **Secondary school or more (in %)** | **p-value from two-sided t-test** |
| Hypertension | Fulfills diagnostic testing criteria | 100·0 | 100·0 | 100·0 | |
| | Tested | 59·5 (58·5-60·5) | 61·9 (61·1-62·8) | 62·6 (61·6-63·6) | <0·0001* |
| | Tested out of all fulfilling criteria | 59·5 (58·5-60·5) | 61·9 (61·1-62·8) | 62·6 (61·6-63·6) | <0·0001* |
| | Guidelines adhered | 59·5 (58·5-60·5) | 61·9 (61·1-62·8) | 62·6 (61·6-63·6) | <0·0001* |

Note: 95% confidence intervals in brackets. The p-values were calculated using an immediate form of a two-sample t-test with unequal variances. We compared male vs. female sex, less than primary vs. secondary or more education, and richest vs. remaining four wealth quintiles. Statistical significance levels based on Benjamini-Hochberg adjusted p-values corresponding to a 5% significance level represented with *.

## 6.3. Appendix for Essay 3

# Appendix 3.1: Literature on the impact of patent expiry on drug consumption

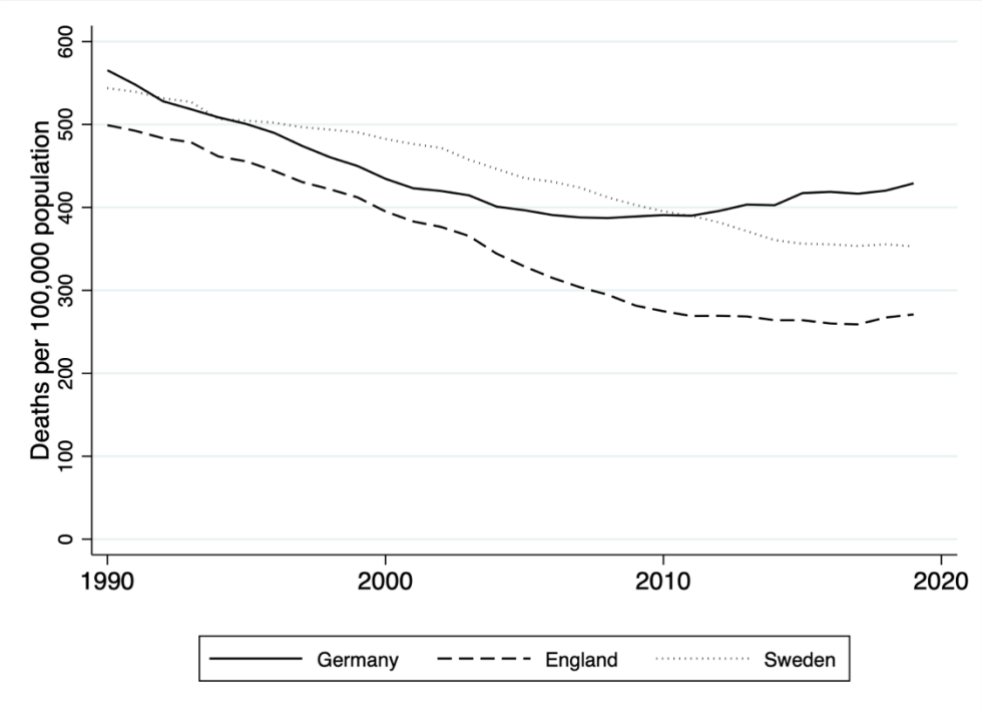
*Appendix table A3.16: Overview of studies examining the impact of patent expiry on a molecule's overall consumption*

| Paper | Country | Time period examined | Data frequency | Econometric method | Consumption of molecule post patent expiry |
|---|---|---|---|---|---|
| Aitken et al. (2013) | USA | 06/2009-05/2013 | Monthly | Descriptive | Increased for 4/6 molecules |
| Berndt, Kyle and Ling (2003) | USA | 01/1998-06/1999 | Monthly | Descriptive | Increased for 1 molecule, unchanged for 1 molecule |
| Chapman, Fitzpatrick and Aladul (2017) | England | 1998-2015 | Yearly | Interrupted Time Series | Increased |
| Duflos and Lichtenberg (2012) | USA | 2000-2004 | Monthly | Weighted Least Squares with fixed effects | Unchanged |
| Fiorentini, Bruni and Mammi (2022) | Emilia-Romagna region, Italy | 2005-2017 | Monthly | Interrupted Time Series | Increased |
| Imai, Fushimi and Sundell (2018) | Sweden, Japan | 2002-2012/13 | Monthly | Time series (ARIMA) | Increased for 1, decreased for 1, unchanged for 16 molecules |
| Lakdawalla and Philipson (2012) | USA | 1990-2003 | Quarterly | Time series with fixed effects | Decreased in short run, increased in long run |

Note: Please refer to the appendix references for full references.

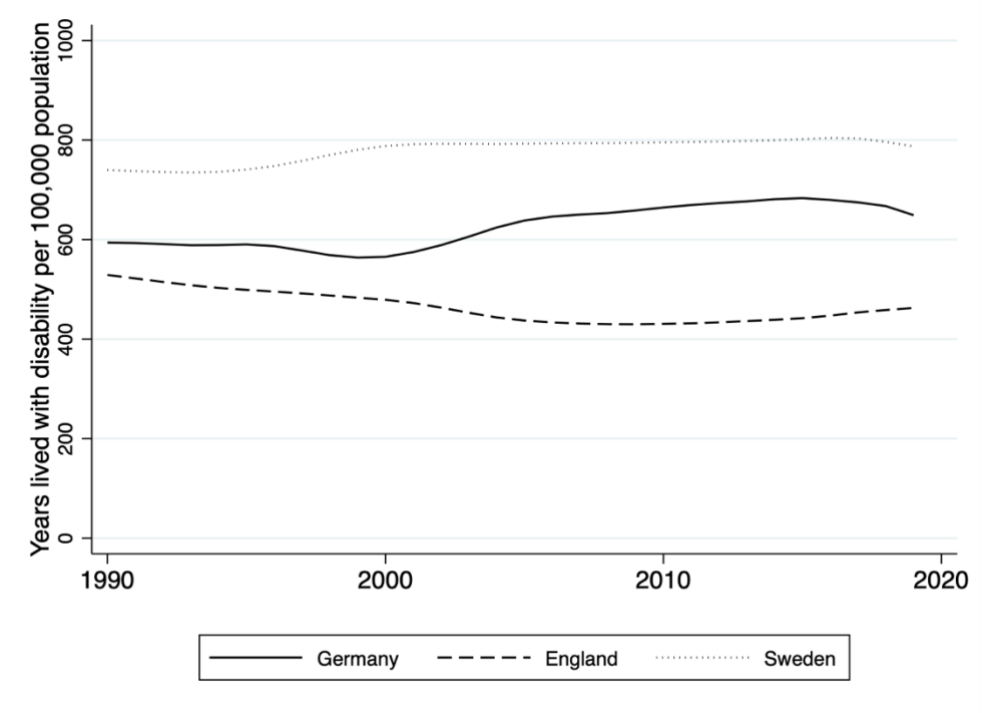
## Appendix 3.2: Cardiovascular disease burden in Germany, England and Sweden

*Appendix figure A3.14: Deaths caused by cardiovascular disease in Germany, England and Sweden*




Source: IHME (2023)

*Appendix figure A3.15: Years lived with disability due to cardiovascular disease in Germany, England and Sweden*




Source: IHME (2023)

## Appendix 3.3: Germany's reference price system

A reference price has to fulfill three conditions: First, the reference price has to be in the bottom 33% of the interval of the lowest and highest price of a standard package. Second, at least 20% of all prescriptions and 20% of all packages must have a price below the reference price. Third, after disregarding outliers, a maximum of 160 prescriptions and packages should be more expensive than the reference price (§35 section 5 German Social Code (SGB) Fifth Book).

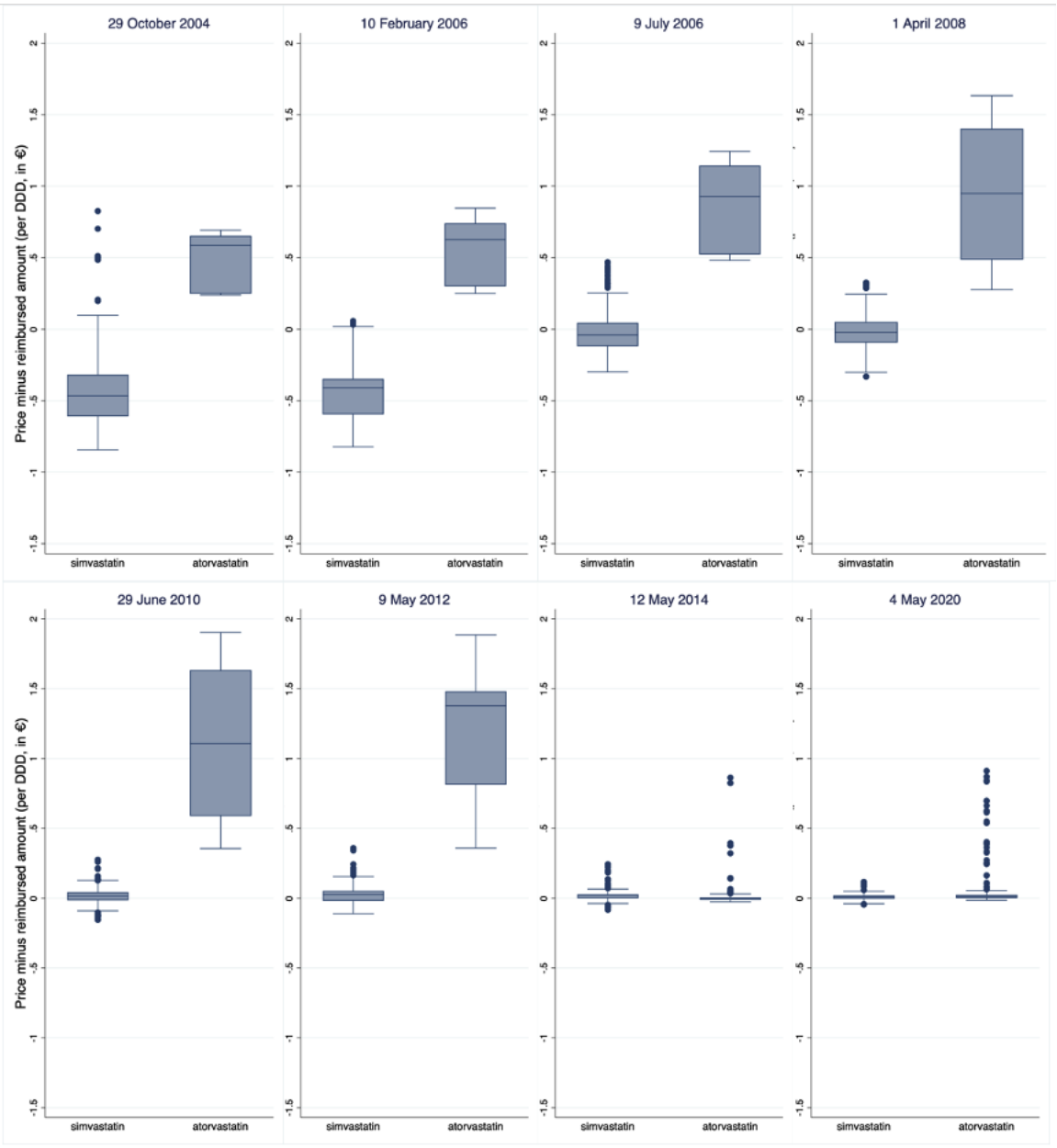
In detail, reference prices are set for a so-called "standard package" with a set strength and number of pills so that an additional calculation is required to determine the reimbursement amount for packages of a specific molecule, dosage and number of pills. All statins were grouped together in a group called "HMG-CoA-Reduktasehemmer" with the introduction of reference prices for therapeutically similar drugs in January 2005. Since then, the number of pills for a standard package in this group has been 100 pills. Appendix table A3.17 below shows that reference prices have been adjusted downwards multiple times in 2006, 2008, 2010, 2012, 2014 and 2020.

*Appendix table A3.17: Reference prices for the group of statins*

| **Date from which reference price was valid (DD/MM/YYYY)** | **Date until which reference price was valid (DD/MM/YYYY)** | **Strength factor** | **Reference price (in €)** | **Reference** |
|---|---|---|---|---|
| 01/05/2005 | 30/03/2006 | 0.97 | 62.55 | GKV-Spitzenverband (2004) |
| 01/04/2006 | 30/06/2006 | 0.97 | 59.42 | GKV-Spitzenverband (2006a) |
| 01/07/2006 | 31/05/2008 | 0.97 | 36.61 | GKV-Spitzenverband (2006b) |
| 01/06/2008 | 31/08/2010 | 0.4 | 13.48 | GKV-Spitzenverband (2008) |
| 01/09/2010 | 30/06/2012 | 0.7 | 11.63 | GKV-Spitzenverband (2010) |
| 01/07/2012 | 30/06/2014 | 0.7 | 7.08 | GKV-Spitzenverband (2012) |
| 01/07/2014 | 31/12/2019 | 0.7 | 5.57 | GKV-Spitzenverband (2014) |
| 01/01/2020 | 30/06/2020 | 0.7 | ? | GKV-Spitzenverband (2019) |
| 01/07/2020 | Current | 0.7 | 4.45 | GKV-Spitzenverband (2020) |

We illustrate the difference between products' reimbursed amount their price in appendix figure A3.16. We calculated the price and reimbursed amount per DDD for each package, ie each molecule-dosage-number of pills combination. These price data were obtained from the years in which the National Association of Statutory Health Insurance Funds announced changes to the reference price of the lipid-modifying agent group (citations in appendix table A3.17) but should be taken as snapshots of the prices at the time of the reference price change decisions as pharmaceutical companies can change prices twice per month, ie we do not observe the prices in between the dates displayed in appendix figure A3.16.

*Appendix figure A3.16: Product prices vis-à-vis reimbursed amount at individual points in time*




Note: We removed outliers where the difference between price and reimbursed amount was more than two standard deviations below or above the mean difference between price and reimbursed amount.

To derive the reference price for a specific product, some calculations using a so-called regression equation and the addition of wholesalers and pharmacies' fees and VAT are required. The regression equation output (explained below with an example) is multiplied with the price of a standard package and thereby accounts for products having varying molecules, dosages and number of pills. Lastly, the German regulation on the prices of medicines ("ArzneimittelPreisverordnung (AMPreisV)") determines that wholesalers can claim a maximum of 3,15% per package (but no more than 37,80€) plus a fixed amount of 0,70€ per package for the buying, storing and distribution of drugs. Pharmacies are allowed to add 3% of the purchase price plus 8,35€ per package plus 0,21€ for the facilitation of emergency service shifts. The reimbursed amount of any statin product can be calculated as follows:

$$Reimbursed\ amount = (rp_{sp} * p + fees) * VAT$$

(13)

Where $rp_{sp}$ refers to the reference price of the standard package, $p$ is the regression equation result and $fees$ can be calculated as follows:

$$fees = 0.0315 * rp_{sp} + 0.7 + 0.03 * rp_{sp} + 8.35 + 0.21$$

(14)

For illustration, I explain this process with an example product and the reference price set for July 2020. This is when a reference price of 4.45€ for a standard package with a strength factor of 0.97 and 100 pills was set. In other words, $rp_{sp} = 4.45$. The fees for our example package would therefore amount to:

$$fees = 0.0315 * 4.45 + 0.7 + 0.03 * 4.45 + 8.35 + 0.21 \approx 9.53368$$

(15)

The regression equation output $p$ is given together with changes to the reference prices (GKV-Spitzenverband) and reads as follows:

$$p = 0.018078546 * wvg^{0.882879} * pk^{0.939798}$$

(16)

where $pk$ refers to the package size and $wvg$ ("Wirkstärkenvergleichsgröße" in German) is the quotient of the package strength divided by the comparison size. The comparison size can be calculated but is also provided with changes to the reference prices for all packages and should be rounded to one decimal digit. The value of the variable "comparison size" is assigned to each molecule whenever the reference price is changed and corresponds to the average amount of milligrams per pill prescribed to patients. Simvastatin and atorvastatin were both assigned a comparison size of 28.7, ie the average simvastatin or atorvastatin pill prescribed contain a dosage of 28.7 mg.

The example product is a package of 30 pills, each containing 10 mg atorvastatin by the producer HEXAL.[42] In other words, $pk = 30$ and the provided $wvg = 0.3$. The output of the regression equation is therefore:

$$p = 0.018078546 * 0.3^{0.882879} * 30^{0.939798} \approx 0.15266$$

(17)

The reimbursed amount would for our example package with a value added tax of 19% would therefore amount to:

[42] The ID using the German product identification system, the "Pharmazentralnummer (PZN)", of this product is 9122555.

$$Reimbursed\ amount \approx (4.45 * 0.15266 + 9.53368) * 1.19 \approx 12.15 \quad (18)$$

The price of this package is listed as 12.08€, so that patients insured with a public health insurance fund would not have to pay anything except for the standard co-payment of 5€.

## Appendix 3.4: Germany's preferred substance(s) and the corresponding target expenditure shares

Since an amendment to §84 of the German Social Code Book V came into effect in May 2006 (buzer, 2023), the National Association of Statutory Health Insurance Physicians ("Kassenärztliche Bundesvereinigung", KBV) sets expenditure targets for drug classes high in consumption volumes in the following way. It groups together molecules within drug classes – in our case statins – and determines one or more preferred molecule(s) that should make up a certain percentage of the total drug class's expenditure. This is what is shown in Table A3.18 below from 2009 onwards.

In 2007 and 2008, simvastatin was also the preferred substance already but there were no national target expenditure shares defined yet. Instead, in 2007 the KBV determined improvement targets for each regional Associations of Statutory Health Insurance Physicians which should reduce the difference between their own and the third best current expenditure share of all regional associations by a third (Kassenärztliche Bundesvereinigung, 2007). In 2008, the KBV determined improvement targets for each regional association in a similar way, only this time recommending to reduce the difference between their own the best current expenditure of all regional associations by a third (Kassenärztliche Bundesvereinigung, 2008a).

*Appendix table A3.18: The preferred substance(s) and the corresponding target percentages of total statin expenditure*

| Year | Preferred substance(s) | Target expenditure share (in %) | Reference |
|---|---|---|---|
| 2009 | simvastatin | 85.0 | Kassenärztliche Bundesvereinigung, 2008b |
| 2010 | simvastatin | 89.0 | Kassenärztliche Bundesvereinigung (2009) |
| 2011 | simvastatin | 86.0 | Kassenärztliche Bundesvereinigung (2010) |
| 2012 | simvastatin | 86.0 | Kassenärztliche Bundesvereinigung (2011) |
| 2013 | simvastatin, pravastatin | 93.0 | Kassenärztliche Bundesvereinigung (2012) |
| 2014 | simvastatin, pravastatin | 87.0 | Kassenärztliche Bundesvereinigung (2013) |
| 2015 | simvastatin, pravastatin | 82.0 | Kassenärztliche Bundesvereinigung (2014) |
| 2016 | simvastatin, pravastatin | 77.0 | Kassenärztliche Bundesvereinigung (2015) |
| 2017 | simvastatin, pravastatin | 70.7 | Kassenärztliche Bundesvereinigung (2016) |
| 2018 | simvastatin, pravastatin | 64.7 | Kassenärztliche Bundesvereinigung (2017) |
| 2019 | simvastatin, pravastatin | 58.3 | Kassenärztliche Bundesvereinigung (2018) |
| 2020 | simvastatin, pravastatin | 51.5 | Kassenärztliche Bundesvereinigung (2019) |
| 2021 | simvastatin, pravastatin | 44.9 | Kassenärztliche Bundesvereinigung (2020) |
| 2022 | simvastatin, pravastatin, atorvastatin | 89.2 | Kassenärztliche Bundesvereinigung (2021) |
| 2023 | simvastatin, pravastatin, atorvastatin | 85.2 | Kassenärztliche Bundesvereinigung (2022) |

Note: There were no national but only regional targets provided from 2017 onwards which is why we then show an unweighted mean of all regional targets (own calculations).

## Appendix 3.5: Data sources and cleaning of Swedish price data

We obtained Swedish price data of atorvastatin and simvastatin by downloading an Excel file of the search results from TLV (2023a; 2023b).

The variables of interest that we use are:

| | |
|---|---|
| Varun | an ID variable for each package |
| Gäller fr.o.m. | the starting date of the price's validity |
| Nytt AUP | the price per pill that a patient pays at the pharmacy |
| Styrka | the dosage of the pill in mg |

To simplify, we do not consider the starting day but the starting month of the price's validity. We generate a price time series for each package (using the Varun ID variable) and then calculate the average price per DDD per month.

## Appendix 3.6: Data sources on lipid-modifying agent consumption

We collected data on a country-by-country basis, searching for publicly available data and requesting data from relevant authorities when no public data was available. We use data provided directly from authorities, such as health ministries, or research arms of health insurance organizations. The yearly data from the US is an exception as we use the yearly, nationally representative Medical Expenditure Panel Survey (MEPS) instead.

*Appendix table A3.19: The availability and sources of the used data*

| Country | Time frame | Availability | Source |
|---|---|---|---|
| **Monthly** | | | |
| England | 02/2008 – 12/2019 | Public | NHSBSA (National Health Service Business Services Authority) |
| Germany | 01/2000 – 12/2021 | Requested | German Institute for Drug Use Evaluation (DAPI) |
| Sweden | 01/2006-10/2021 | Public | Socialstyrelsen (National Board of Health and Welfare, a government agency of the Ministry of Health) |
| **Yearly** | | | |
| Belgium | 1997-2020 | Requested | Farmanet database of RIZIV (National Institute for Health and Disability Insurance) |
| Croatia | 2004-2020 | Requested | HALMED (Agency for Medicinal Products and Medical Devices) |
| Denmark | 1997-2020 | Public | Danish Health Data Authority (2023) |
| Estonia | 2010-2018 | Public | Baltic Statistics on Medicines[1] |
| Finland | 2002-2021 | Requested | FIMEA (Finnish Medicines Agency) |
| Latvia | 2010-2018 | Public | Baltic Statistics on Medicines[1] |
| Lithuania | 2010-2018 | Public | Baltic Statistics on Medicines[1] |
| Malaysia | 2004-2016 | Public | Malaysian Statistics on Medicines[2] |
| Netherlands | 2003-2020 | Requested | Zorginstituut Nederland |
| Norway | 2004-2020 | Public | Norwegian Institute of Public Health (2021) |
| Scotland | 2001-2016 | Public | Public Health Scotland (2020) |
| Spain | 2009-2020 | Requested | Ministry of Health (Subdirección de Farmacia) |
| USA | 1996-2019 | Public | Medical Expenditure Panel Survey (MEPS) from the Agency for Healthcare Research and Quality (2023) |

Note: Please refer to the appendix references for the publicly available references.

[1] Estonia, Latvia and Lithuania's lipid-modifying agent consumption was taken out of three reports entitled Baltic Statistics on Medicines published by the Estonian State Agency of Medicines (2013), the Latvian State Agency of Medicines (2016) and the Lithuanian State Agency of Medicines (2019).

[2] The Malaysian Statistics on Medicines were annual reports until 2008 (Ministry of Health Malaysia, 2006; 2007; 2009; 2010; 2013) and then for several years jointly thereafter (Ministry of Health Malaysia, 2014; 2017; 2020).

## Appendix 3.7: Explanation converting English data to DDD

The United Kingdom uses its own system to classify and group drugs called the British National Formulary (BNF) based on 15 digit long codes (the first seven digits correspond to the category of the drug and the last eight digits represent the medicinal product, form and strength), with descending hierarchy levels called chapters, sections, paragraphs, chemical substances and individual preparations (see the Glossary in NHSBSA, 2021a).

Prescription cost analysis data from the National Health Service Business Services Authority (NHSBSA) were provided monthly from February 2008 to March 2021 (NHSBSA, 2021a). This data records all drugs dispensed in the community in England, based on the prescriptions issued by GPs and other authorized prescribers. This means that prescriptions issued in hospitals but dispensed in community pharmacies are included while those dispensed in the hospital pharmacies are not included in this data. The data contain the prescriptions dispensed by community pharmacies (the majority) as well as dispensing doctors (used less, for instance, in rural areas). The used and described procedures are similar to those used by Chapman, Fitzpatrick and Aladul (2017).

**Lipid-lowering (C10) molecules**
For each month, we used all data points where the BNF section name read “Lipid-Regulating Drugs” and appended the data from the pharmacy and doctor dispensers. We used the mg quantities indicated in the ‘drug name’ variable and the ‘quantity’ variable to determine the total milligram quantity consumed for each package-strength combination. We then sum the milligram quantities of each package-strength combination over each molecule, assign the corresponding ATC code, and convert the milligram total to DDD using the ATC definitions (WHOCC, 2018).

Example data row:

| BNF Section Name | BNF Chemical Name | Drug Name | Quantity |
|---|---|---|---|
| Lipid-Regulating Drugs | Simvastatin | Simvastatin_Tab 10mg | 6551208 |

**Control molecules (ATC groups C02, C03, C07, C08, C09)**
The same procedure was followed as the C10 molecules except for the first step searching for all relevant package-strength rows. Here, a combination of searching for BNF section names such as “Hypertension and Heart Failure” or “Nit, Calc Block & Other Antianginal Drugs” and individual molecule names such as “Felodipine” was used. We ensured that all molecules contained in the relevant BNF sections were indeed part of one of the ATC level 2 groups we selected. For instance, the molecule “Isosorbide Mononitrate” was included in the “Nit, Calc Block & Other Antianginal Drugs” group but due to its classification as an ATC-group C01 molecule, we did not include it in our “to-DDD-conversion” exercise.

## Appendix 3.8: The impact of atorvastatin's patent expiry on simvastatin consumption

Here, we show the results of an interrupted time series analysis. Atorvastatin's patent expiry in May 2012 negatively impacted simvastatin consumption in all three countries.

*Appendix figure A3.17: The impact of atorvastatin's patent expiry on simvastatin consumption using interrupted time series*


Note: The vertical red lines correspond to the date of the atorvastatin's patent expiry.

*Appendix table A3.20: Interrupted time series results of the impact of atorvastatin's patent expiry on simvastatin's consumption*

| | **Germany (1)** | **England (2)** | **Sweden (3)** |
|---|---|---|---|
| time ($\beta_1$) | 8.13*** | 8.71*** | 0.050*** |
| | (0.665) | (1.18) | (0.019) |
| expiry ($\beta_2$) | -56.5** | -78.5 | -0.908** |
| | (24.3) | (47.4) | (0.383) |
| time post expiry ($\beta_3$) | -10.2*** | -28.4*** | -0.192*** |
| | (1.07) | (2.00) | (0.021) |
| constant ($\beta_0$) | 1328*** | 2274*** | 24.5*** |
| | (16.7) | (30.8) | (0.520) |
| Observations | 80 | 80 | 80 |

Note: We employ Newey-West standard errors correcting for heteroscedasticity and autocorrelation and display statistical significance levels (10%, 5%, 1%) with */**/***.

## Appendix 3.9: Patent protection status of donor pool molecules

*Appendix table A3.21: Patent protection status of the control donor pool molecules for the SCM analyses*

| ATC | expiry | ATC | expiry | ATC | expiry | ATC | expiry | ATC | expiry |
|---|---|---|---|---|---|---|---|---|---|
| C01AA04 | Before 2000 | C02AC02 | Before 2000 | C03DB01 | Before 2000 | C07AB05 | Before 2000 | C09AA16 | ? |
| C01AA05 | Before 2000 | C02AC05 | 2003 | C03DB02 | Before 2000 | C07AB07 | Before 2000 | C09BA01 | Before 2000 |
| C01BA01 | Before 2000 | C02CA01 | Before 2000 | C03EA01 | Before 2000 | C07AB08 | Before 2000 | C09BA02 | Before 2000 |
| C01BA03 | ? | C02CA02 | ? | C03EA06 | ? | C07AB09 | Before 2000 | C09BA03 | Before 2000 |
| C01BB01 | Before 2000 | C02CA04 | Before 2000 | C03EA07 | ? | C07AB12 | Before 2000 | C09BA04 | Before 2000 |
| C01BB02 | ? | C02CA06 | 2016 | C03EB01 | ? | C07AG01 | Before 2000 | C09BA05 | Before 2000 |
| C01BC03 | Before 2000 | C02CC02 | ? | C03EB02 | ? | C07AG02 | Before 2000 | C09BA06 | Before 2000 |
| C01BC04 | 2003 | C02DA01 | ? | C03XA01 | Not expired | C07FB02 | Not expired | C09BA08 | Before 2000 |
| C01BD01 | Before 2000 | C02DB01 | Not expired | C04AB01 | ? | C08CA01 | 2004 | C09BB05 | Not expired |
| C01BD05 | ? | C02DB02 | ? | C04AD03 | Before 2000 | C08CA02 | Before 2000 | C09BB06 | Not expired |
| C01BD07 | Not expired | C02DC01 | Not expired | C04AE01 | Before 2000 | C08CA03 | ? | C09BB10 | Not expired |
| C01BG07 | ? | C02KD01 | ? | C04AX02 | Not expired | C08CA04 | Before 2000 | C09BX01 | Not expired |
| C01CA01 | Before 2000 | C02KX01 | 2017 | C04AX21 | Before 2000 | C08CA05 | Before 2000 | C09CA01 | 2010 |
| C01CA02 | ? | C02KX02 | Not expired | C05AA01 | Before 2000 | C08CA06 | 2001 | C09CA02 | 2008 |
| C01CA03 | ? | C02KX03 | Removed in 2010 | C05AA04 | Before 2000 | C08CA07 | Before 2000 | C09CA03 | 2005 |
| C01CA04 | ? | C02KX04 | Not expired | C05AA08 | Before 2000 | C08CA08 | Before 2000 | C09CA04 | 2012 |
| C01CA06 | Before 2000 | C02KX05 | Not expired | C05AE01 | Before 2000 | C08CA09 | Before 2000 | C09CA06 | 2012 |
| C01CA07 | ? | C02LA01 | ? | C05AX03 | NA | C08CA10 | Before 2000 | C09CA07 | 2013 |
| C01CA17 | Not expired | C03AA01 | Not expired | C05BA01 | ? | C08CA11 | ? | C09CA08 | 2017 |
| C01CA24 | Before 2000 | C03AA03 | Before 2000 | C05BA04 | ? | C08CA13 | 2010 | C09CA09 | ? |
| C01CX08 | Not expired | C03AA04 | ? | C05BB02 | Before 2000 | C08DA01 | Before 2000 | C09DA01 | 2010 |
| C01DA02 | Before 2000 | C03AA05 | ? | C05BB04 | ? | C08DA51 | ? | C09DA02 | 2008 |
| C01DA08 | Before 2000 | C03AA07 | ? | C05CA01 | ? | C08DB01 | Before 2000 | C09DA03 | 2003 |
| C01DA14 | Before 2000 | C03AB01 | ? | C05CA03 | ? | C08GA | ? | C09DA04 | 2014 |
| C01DX12 | Before 2000 | C03BA04 | Not expired | C05CA04 | Removed in 2002 | C08GA01 | ? | C09DA06 | 2012 |
| C01DX16 | ? | C03BA08 | ? | C05CA53 | ? | C09AA01 | Before 2000 | C09DA07 | 2014 |
| C01DX22 | Not expired | C03BA10 | 2004 | C07AA02 | Before 2000 | C09AA02 | Before 2000 | C09DA08 | 2018 |
| C01EA01 | ? | C03BA11 | Before 2000 | C07AA03 | Before 2000 | C09AA03 | Before 2000 | C09DB01 | 2019 |
| C01EB09 | ? | C03CA01 | Before 2000 | C07AA05 | Before 2000 | C09AA04 | Before 2000 | C09DB02 | 2019 |
| C01EB10 | ? | C03CA02 | ? | C07AA06 | Before 2000 | C09AA05 | Before 2000 | C09DX03 | Not expired |
| C01EB15 | ? | C03CA04 | 2003 | C07AA07 | Before 2000 | C09AA06 | Before 2000 | C09DX04 | Not expired |
| C01EB17 | 2017 | C03CC01 | Before 2000 | C07AA12 | Before 2000 | C09AA08 | Before 2000 | C09XA02 | Not expired |
| C01EB18 | Not expired | C03DA01 | Before 2000 | C07AB02 | Before 2000 | C09AA09 | Before 2000 | | |
| C02AB01 | Before 2000 | C03DA04 | 2014 | C07AB03 | Before 2000 | C09AA10 | Before 2000 | | |
| C02AC01 | Before 2000 | C03DA05 | Not expired | C07AB04 | ? | C09AA13 | ? | | |

Source: Klose and Schwabe, 1999; 2001; 2002; 2003; 2004a; 2004b; 2006; 2007; 2008a; 2008b; 2009; 2010; 2011; 2012; 2013; 2014; 2015; 2016; 2017; 2018; 2019; 2020.

# Appendix 3.10: Medical guidelines

*Appendix table A3.22: German, English, Swedish and select international medical guidelines on lipid-modifying agent consumption*

| Year | Organization | Guideline | Specifically recommends simvastatin? | Specifically recommends atorvastatin? | Reference |
|---|---|---|---|---|---|
| 1998 | Second Joint Task Force of European and other Societies on Coronary Prevention | Prevention of coronary heart disease in clinical practice | No | No | Wood et al. (1998a) |
| 1998 | Joint British Societies | "JBS 1" Joint British recommendations on prevention of coronary heart disease in clinical practice | No | No | Wood et al. (1998b) |
| 1998 | German Cardiac Society (DGK) | Guideline: Coronary Artery Disease / Angina pectoris | No | No | Meyer et al. (1998) |
| 2000 | British Health Ministry | National service framework: coronary heart disease | No | No | Department of Health and Social Care (2000) |
| 2001 | American Heart Association (AHA) and American College of Cardiology (ACC) | Guidelines for Preventing Heart Attack and Death in Patients with Atherosclerotic Cardiovascular Disease: 2001 Update | No | No | Smith et al. (2001) |
| 2001 | German Cardiac Society (DGK) | Recommendations on comprehensive risk reduction for patients with coronary artery disease, angiopathy and diabetes | No | No | Gohlke et al. (2001) |
| 2003 | Third Joint Task Force of European and other Societies on Cardiovascular Disease Prevention in Clinical Practice | European guidelines on cardiovascular disease prevention in clinical practice | No | No | De Backer et al. (2003) |
| 2003 | German Cardiac Society (DGK) | Position paper on primary prevention of cardiovascular diseases | No | No | Gohlke et al. (2003) |
| 2003 | German Cardiac Society (DGK) | Guideline for the diagnosis and treatment of chronic coronary artery disease | No | No | Dietz and Rauch (2003) |
| 2005 | Joint British Societies[4] | "JBS 2" Joint British Societies' guidelines on prevention of cardiovascular disease in clinical practice | No | No | Wood et al. (2005) |
| 2005 | German Cardiac Society (DGK) | Supplementary volume primary prevention of cardiovascular diseases | No | No | Scheller et al. (2005) |
| 2006 | American Heart Association (AHA) and American College of Cardiology (ACC) | Guidelines for Secondary Prevention for Patients with Coronary and Other Atherosclerotic Vascular Disease: 2006 Update | No | No | Smith et al. (2006) |

| | | | | | |
|---|---|---|---|---|---|
| 2006 | German Cardiac Society (DGK) | Recommendations on the diagnostic and treatment of patients with coronary artery disease and renal failure | No | No | Reinecke et al. (2006) |
| 2006 | National Programme for Disease Management Guidelines (NVL) Germany | National Disease Management Guideline for chronic coronary heart disease 1.0 | No | No | Donner-Banzhoff et al. (2006) |
| 2006 | National Institute for Health and Care Excellence (NICE) | Statins for the prevention of cardiovascular events (Technology Appraisal Guidance 94) | No | No | NICE (2006) |
| 2007 | Fourth Joint Task Force of the European Society of Cardiology and other societies[5] | European guidelines on cardiovascular disease prevention in clinical practice: executive summary | No | No | Graham et al. (2007) |
| 2007 | German Cardiac Society (DGK) | Position paper on statin therapy | No | No | Böhm et al. (2007) |
| 2007 | German Cardiac Society (DGK) | Guideline on risk-adjusted prevention of cardiovascular diseases | No | No | Gohlke et al. (2007) |
| 2008 | National Board of Health and Welfare (socialstyrelsen) Sweden | National guidelines for cardiac care 2008 | Yes | No | National Board of Health and Welfare (2008) |
| 2008 | National Institute for Health and Care Excellence (NICE) | Lipid modification (Clinical guideline 67) | Yes | For niche cases* | NICE (2008a) |
| 2008 | National Institute for Health and Care Excellence (NICE) | Cardiovascular disease: identifying and supporting people most at risk of dying early | No | No | NICE (2008b) |
| 2011 | American Heart Association (AHA) | Effectiveness-Based Guidelines for the Prevention of Cardiovascular Disease in Women – 2011 Update | No | No | Mosca et al. (2011) |
| 2011 | American Heart Association (AHA) and American College of Cardiology Foundation (ACCF) | Secondary Prevention and Risk Reduction Therapy for Patients with Coronary and Other Atherosclerotic Vascular Disease: 2011 Update | No | No | Smith et al. (2011) |
| 2011 | Task Force on the Diagnosis and Treatment of Peripheral Artery Diseases of the European Society of Cardiology (ESC) | ESC Guidelines on the diagnosis and treatment of peripheral artery diseases | No | No | Tendera et al. (2011) |
| 2011 | Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and the European Atherosclerosis Society (EAS) | ESC/EAS Guidelines for the management of dyslipidaemias | No | No | Reiner et al. (2011) |
| 2012 | Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice[6] | European Guidelines on cardiovascular disease prevention in clinical practice | No | No | Perk et al. (2012) |

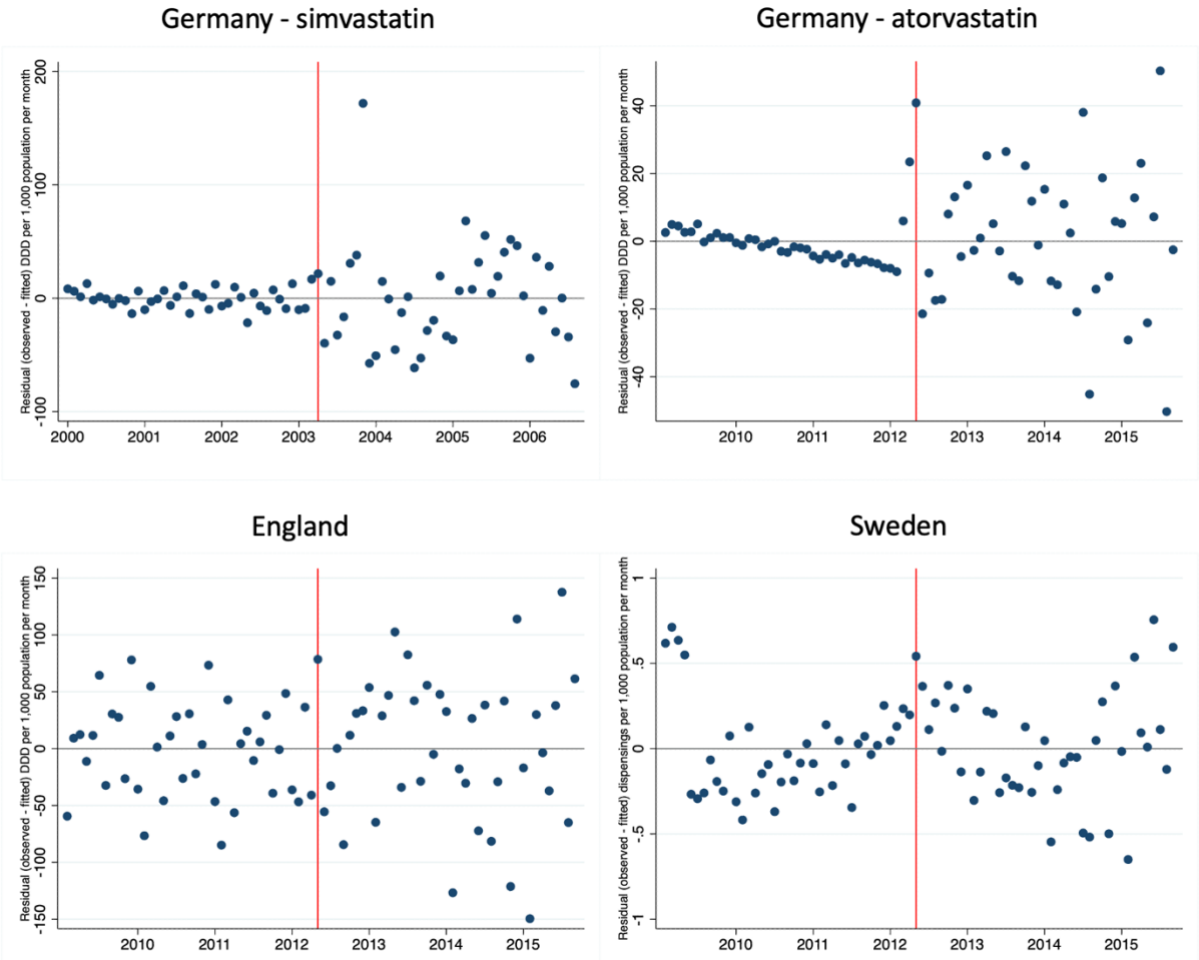
| | | | | | |
|---|---|---|---|---|---|
| 2012 | German Cardiac Society (DGK) | Comment on the 2011 guideline of the European Society of Cardiology on the diagnosis and therapy of peripheral vascular diseases | No | No | Erbel et al. (2012) |
| 2012 | German Cardiac Society (DGK) | Comment on the new 2011 guidelines of the European Society of Cardiology on the management of dyslipidemia | No | No | Koenig et al. (2012) |
| 2013 | American College of Cardiology / American Heart Association Task Force on Practice Guidelines | Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults | No | No | Stone et al. (2013) |
| 2013 | Task Force on diabetes, pre-diabetes, and cardiovascular diseases of the European Society of Cardiology (ESC) and developed in collaboration with the European Association for the Study of Diabetes (EASD). | ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD | No | No | Rydén et al. (2013) |
| 2013 | German Cardiac Society (DGK) | Comment on the 2012 guidelines of the European Society of Cardiology on cardiovascular prevention | No | No | Schuler et al. (2013) |
| 2013 | National Programme for Disease Management Guidelines (NVL) Germany | National Disease Management Guideline for chronic coronary heart disease 2.0 | No | No | Donner-Banzhoff et al. (2013) |
| 2014 | Joint British Societies | "JBS 3" Joint British Societies' consensus recommendations for the prevention of cardiovascular disease | No | No | JBS3 Board (2013) |
| 2014 | German Cardiac Society (DGK) | Comment on the 2013 guidelines of the European Society of Cardiology on diabetes, pre-diabetes and cardiovascular diseases | No | No | Marx et al. (2014) |
| 2014 | German Cardiac Society (DGK) | Position regarding the US guidelines on the reduction of the risk of atherosclerosis through lipid-lowering therapy | No | No | Gohlke et al. (2014) |
| 2014 | National Institute for Health and Care Excellence (NICE) | Clinical guideline CG181: Cardiovascular disease: risk assessment and reduction, including lipid modification | No | Yes | NICE (2014) |
| 2014 | National Programme for Disease Management Guidelines (NVL) Germany | National Disease Management Guideline for chronic coronary heart disease 3.0 | No | No | Donner-Banzhoff et al. (2014) |
| 2015 | National Board of Health and Welfare (socialstyrelsen) Sweden | Heart care – Recommendations, assessments and summary | No | No | National Board of Health and Welfare (2015) |
| 2016 | Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice[7] | European Guidelines on cardiovascular disease prevention in clinical practice | No | No | Piepoli et al. (2016) |

| | | | | | |
|---|---|---|---|---|---|
| 2016 | Task Force for the Management of Dyslipidaemias of the European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS) | ESC/EAS Guidelines for the Management of Dyslipidaemias | No | No | Catapano et al. (2016) |
| 2016 | National Programme for Disease Management Guidelines (NVL) Germany | National Disease Management Guideline for chronic coronary heart disease 4.0 | No | No | Laufs et al. (2016) |
| 2017 | German College of General Practitioners and Family Physicians (DEGAM) | General Practitioners' risk advice on cardiovascular prevention | No | No | Ludt et al. (2017) |
| 2017 | German Cardiac Society (DGK) | Comment on the 2016 guidelines of the European Society of Cardiology and European Atherosclerosis Society on the diagnosis and therapy of dyslipidemias | No | No | Landmesser et al. (2017) |
| 2018 | AHA/ACC/AACVPR/AAPA/ ABC/ACPM/ADA/AGS/ APhA/ASPC/NLA/PCNA[8] | Guideline on the Management of Blood Cholesterol | No | No | Grundy et al. (2018) |
| 2018 | National Board of Health and Welfare (socialstyrelsen) Sweden | National guidelines for cardiac care | No | No | National Board of Health and Welfare (2018) |
| 2019 | Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS) | ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk | No | No | Mach et al. (2020) |
| 2019 | National Programme for Disease Management Guidelines (NVL) Germany | National Disease Management Guideline for chronic coronary heart disease 5.0 | No | No | Dißmann et al. (2019) |
| 2022 | National Programme for Disease Management Guidelines (NVL) Germany | National Disease Management Guideline for chronic coronary heart disease 6.0 | No | No | Schneider et al. (2022) |

Notes:
*They recommend higher intensity statins like 80mg simvastatin or atorvastatin for individuals with acute coronary syndrome.

# Appendix 3.11: Autocorrelation testing

*Appendix figure A3.18: Residual plots of interrupted time series analyses*



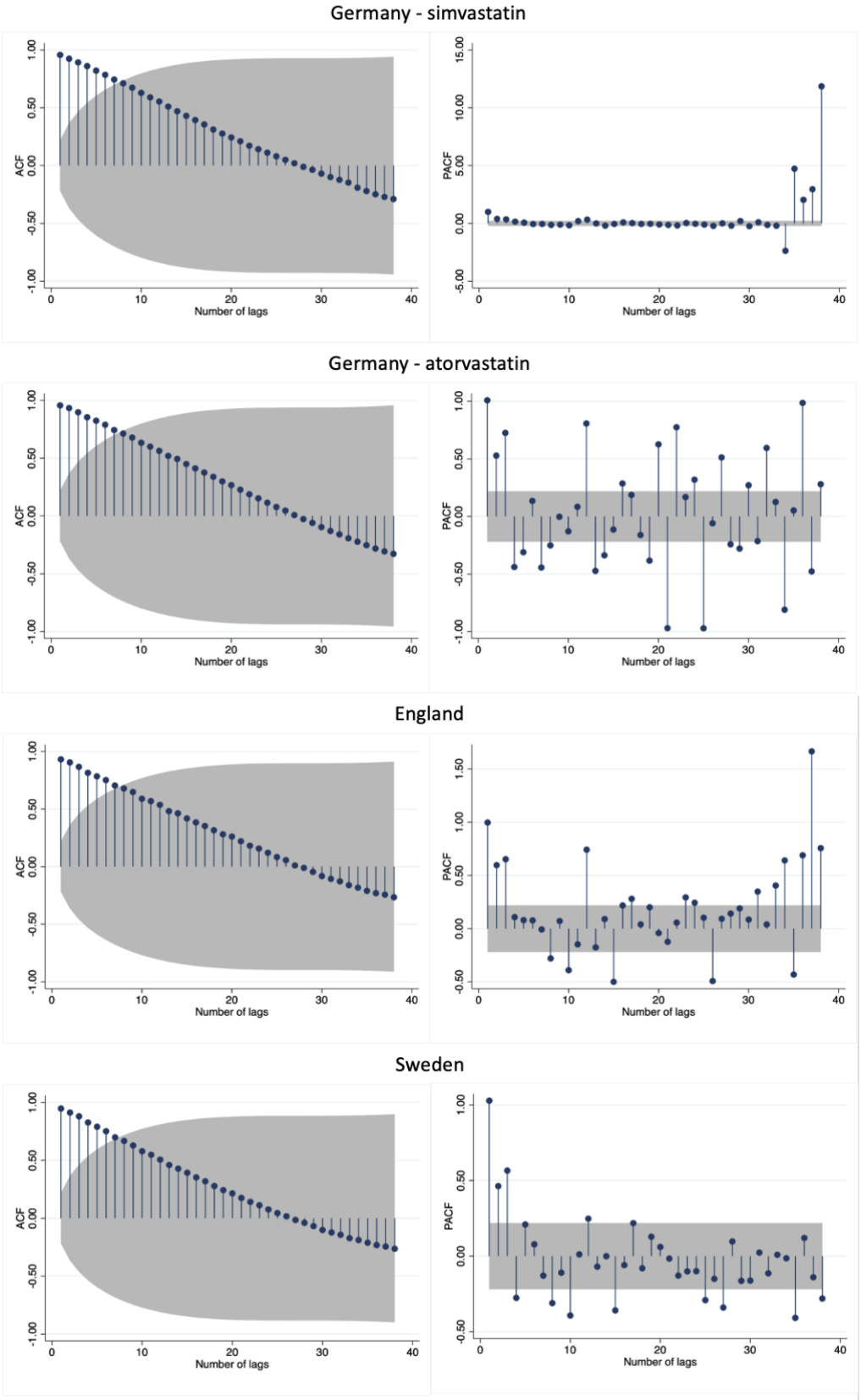

Note: Vertical line corresponds to date of patent expiry.

*Appendix table A3.23: Durbin Watson test statistics and autocorrelation bounds*

| | Durban Watson test statistic | 4 – Durban Watson test statistic | Bounds for 80 observations and 3 regressors excl. the intercept for … | | | | Positive autocorrelation | | Negative autocorrelation | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 1% significance | | 5% significance | | | | | |
| | | | dL (lower bound) | dU (upper bound) | dL (lower bound) | dU (upper bound) | 1% signi-ficance | 5% signif-icance | 1% signi-ficance | 5% signi-ficance |
| Germany simvastatin | 1.678 | 2.322 | 1.416 | 1.568 | 1.560 | 1.715 | No | Inconclusive | No | No |
| Germany atorvastatin | 2.461 | 1.539 | 1.416 | 1.568 | 1.560 | 1.715 | No | No | Inconclusive | Yes |
| **England** | 2.558 | 1.442 | 1.416 | 1.568 | 1.560 | 1.715 | No | No | Inconclusive | Yes |
| **Sweden** | 1.439 | 2.561 | 1.416 | 1.568 | 1.560 | 1.715 | Inconclusive | Yes | No | No |

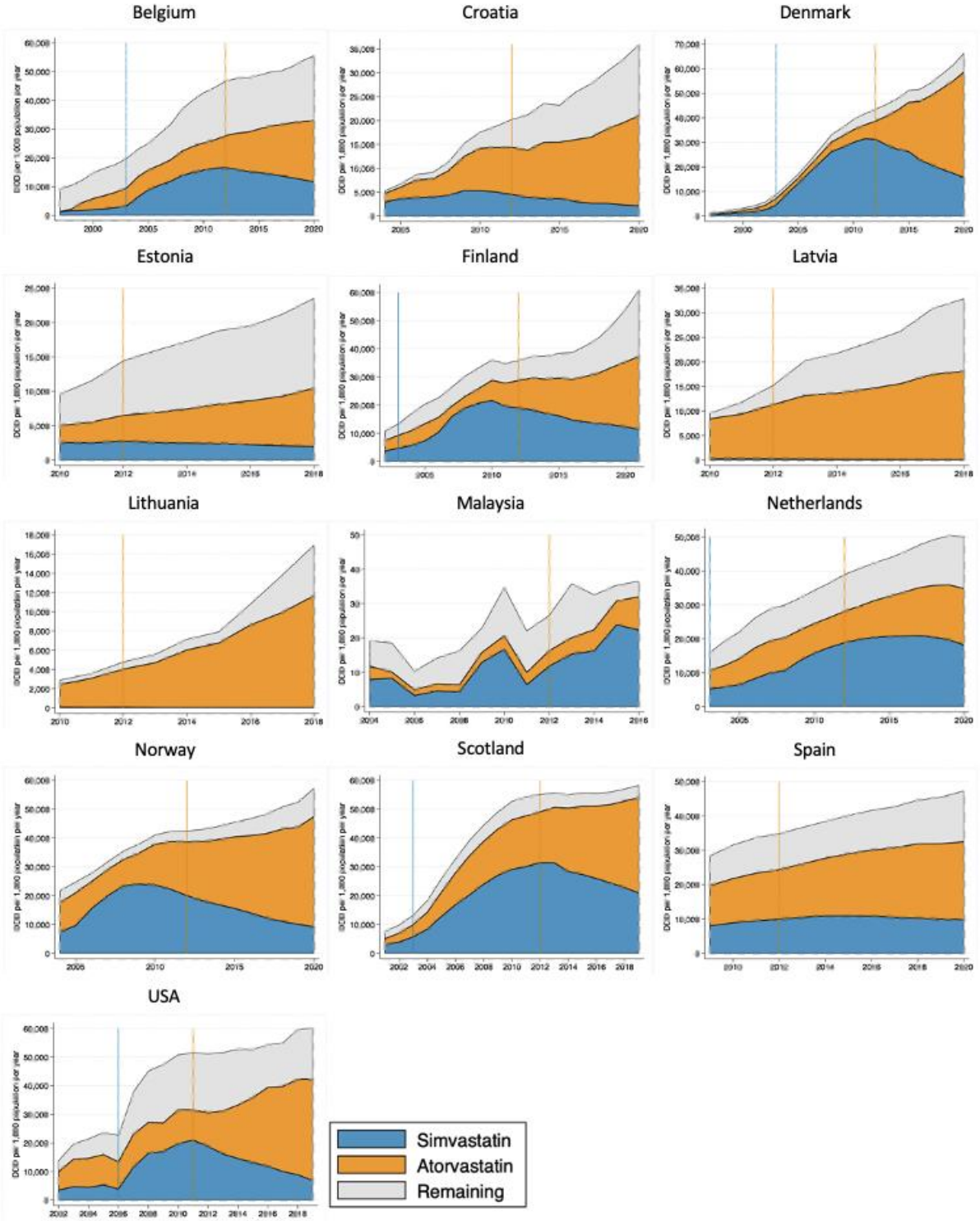
Note: If the Durbin Watson test statistic is lower than dL, we reject the null hypothesis and there is positive autocorrelation. If it is higher than dU, we do not reject the null hypothesis and there is no positive autocorrelation. If it is within the bounds, the test is inconclusive. For negative autocorrelation, you subtract the Durban Watson test statistic from 4 and proceed as with positive autocorrelation. Upper and lower bounds for 1% and 5% significance are taken from Savin and White (1977).

*Appendix figure A3.19: Graphs of the autocorrelation function (ACF) and the partial autocorrelation function (PACF)*



# Appendix 3.12: Yearly lipid-modifying agent consumption

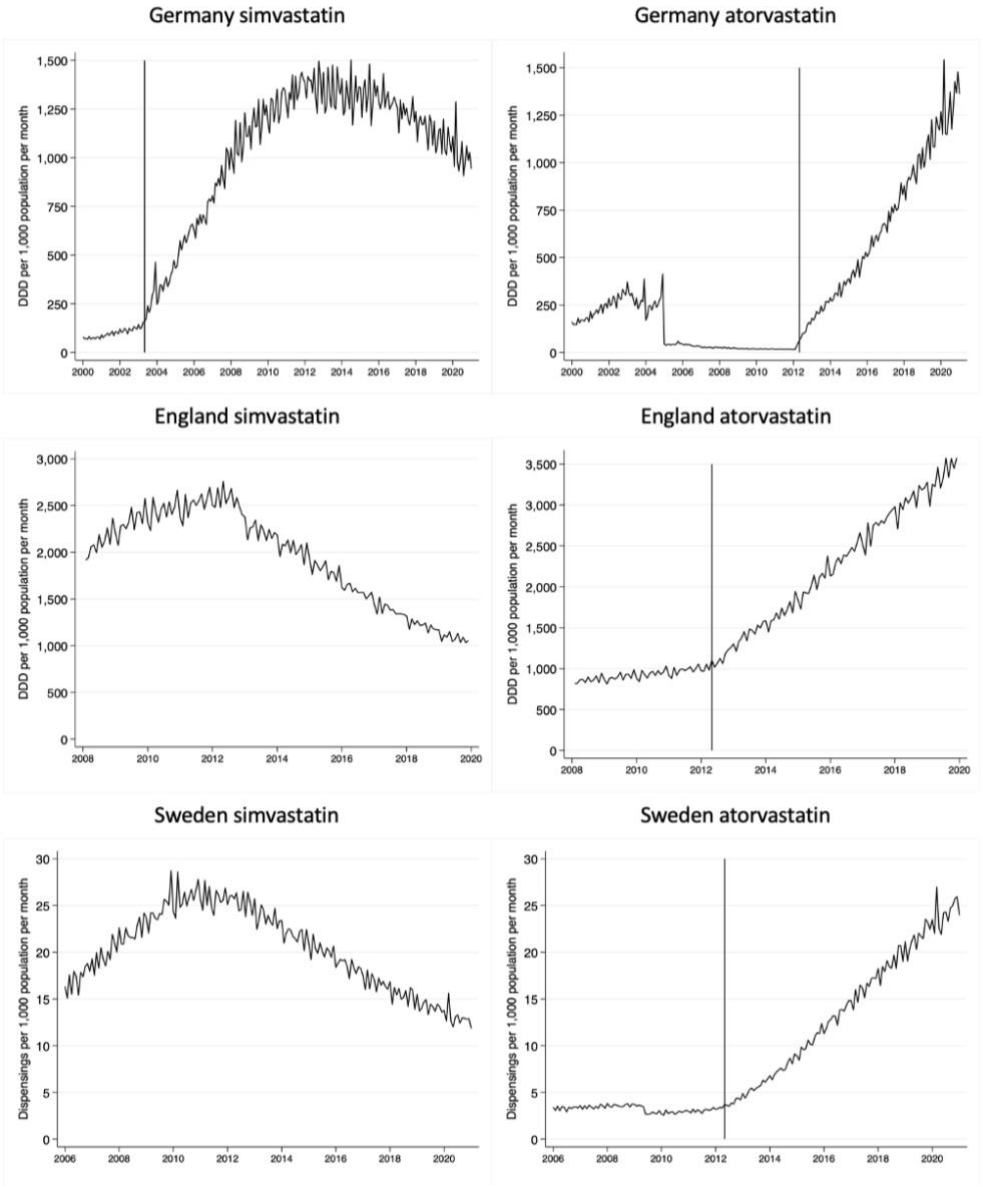
*Appendix figure A3.20: Yearly lipid-modifying agent consumption*




Note: Vertical lines refer to the patent expiry of the molecule with the same color.

# Appendix 3.13: Monthly consumption of simvastatin and atorvastatin

*Appendix figure A3.21: Monthly consumption of simvastatin and atorvastatin in Germany, England and Sweden*




Note: Vertical lines correspond to molecules' patent expiry date.

## Appendix 3.14: Synthetic control weights by country

*Appendix table A3.24: Synthetic control weights of German donor pool molecules*

| ATC | atorvastatin | simvastatin |
|---|---|---|
| C02AB01 | 0.034 | 0.026 |
| C02AC01 | 0.025 | 0.028 |
| C02AC02 | 0.059 | |
| C02AC05 | 0.014 | 0.037 |
| C02CA01 | 0.035 | 0.026 |
| C02CA04 | 0.018 | 0.035 |
| C02CA06 | 0.025 | 0.027 |
| C02DB01 | 0.029 | 0.026 |
| C02DC01 | 0.036 | 0.026 |
| C02KX01 | 0.042 | 0.025 |
| C02KX02 | 0.053 | 0.025 |
| C02KX03 | | 0.025 |
| C02KX04 | 0.059 | |
| C02KX05 | 0.059 | |
| C02LA01 | 0.025 | 0.032 |
| C08CA01 | 0.006 | 0.139 |
| C08CA02 | 0.017 | 0.036 |
| C08CA03 | 0.032 | 0.027 |
| C08CA04 | 0.049 | 0.026 |
| C08CA05 | 0.017 | 0.064 |
| C08CA06 | 0.046 | 0.026 |
| C08CA07 | 0.029 | 0.027 |
| C08CA08 | 0.015 | 0.05 |
| C08CA09 | 0.043 | 0.027 |
| C08CA10 | 0.028 | 0.027 |
| C08CA11 | 0.038 | 0.025 |
| C08CA13 | 0.013 | 0.029 |
| C08DA01 | 0.015 | 0.054 |
| C08DA51 | 0.039 | 0.026 |
| C08DB01 | 0.024 | 0.03 |
| C08GA | 0.031 | 0.026 |
| C08GA01 | 0.045 | 0.026 |

*Appendix table A3.25: Synthetic control weights of English donor pool molecules*

| ATC | Weight | ATC | Weight | ATC | Weight | ATC | Weight |
|---|---|---|---|---|---|---|---|
| C02AB01 | 0.008 | C03BA08 | 0.008 | C08CA07 | 0.008 | C09BB05 | 0.008 |
| C02AC01 | 0.008 | C03BA10 | 0.008 | C08CA09 | 0.008 | C09BB10 | 0.008 |
| C02AC02 | 0.008 | C03BA11 | 0.008 | C08CA13 | 0.009 | C09BX01 | 0.008 |
| C02AC05 | 0.008 | C03CA01 | 0.019 | C08DA01 | 0.008 | C09CA01 | 0.011 |
| C02CA01 | 0.008 | C03CA02 | 0.009 | C08DB01 | 0.009 | C09CA02 | 0.008 |
| C02CA02 | 0.008 | C03CA04 | 0.008 | C09AA01 | 0.008 | C09CA03 | 0.009 |
| C02CA04 | 0.013 | C03DA01 | 0.008 | C09AA02 | 0.011 | C09CA04 | 0.010 |
| C02CC02 | 0.008 | C03DA04 | 0.008 | C09AA03 | 0.018 | C09CA06 | 0.013 |
| C02DA01 | 0.008 | C03DB01 | 0.008 | C09AA04 | 0.012 | C09CA07 | 0.008 |
| C02DB02 | 0.008 | C03DB02 | 0.008 | C09AA05 | 0.265 | C09CA08 | 0.008 |
| C02DC01 | 0.008 | C03EA01 | 0.008 | C09AA06 | 0.008 | C09CA09 | 0.008 |
| C02KD01 | 0.008 | C03EA06 | 0.008 | C09AA08 | 0.008 | C09DA01 | 0.008 |
| C02KX01 | 0.008 | C03EA07 | 0.008 | C09AA09 | 0.008 | C09DA03 | 0.008 |
| C02KX02 | 0.008 | C03EB01 | 0.009 | C09AA10 | 0.008 | C09DA04 | 0.008 |
| C02KX03 | 0.008 | C03EB02 | 0.008 | C09AA13 | 0.008 | C09DA07 | 0.008 |
| C03AA01 | 0.029 | C08CA01 | 0.038 | C09AA16 | 0.008 | C09DA08 | 0.008 |
| C03AA03 | 0.008 | C08CA02 | 0.011 | C09BA01 | 0.008 | C09DB01 | 0.008 |
| C03AA04 | 0.008 | C08CA03 | 0.008 | C09BA02 | 0.008 | C09DB02 | 0.008 |
| C03AA05 | 0.008 | C08CA04 | 0.008 | C09BA03 | 0.008 | C09DX03 | 0.008 |
| C03AA07 | 0.008 | C08CA05 | 0.010 | C09BA04 | 0.008 | C09XA02 | 0.008 |
| C03BA04 | 0.008 | C08CA06 | 0.008 | C09BA06 | 0.008 | | |

*Appendix table A3.26: Synthetic control weights of Swedish donor pool molecules*

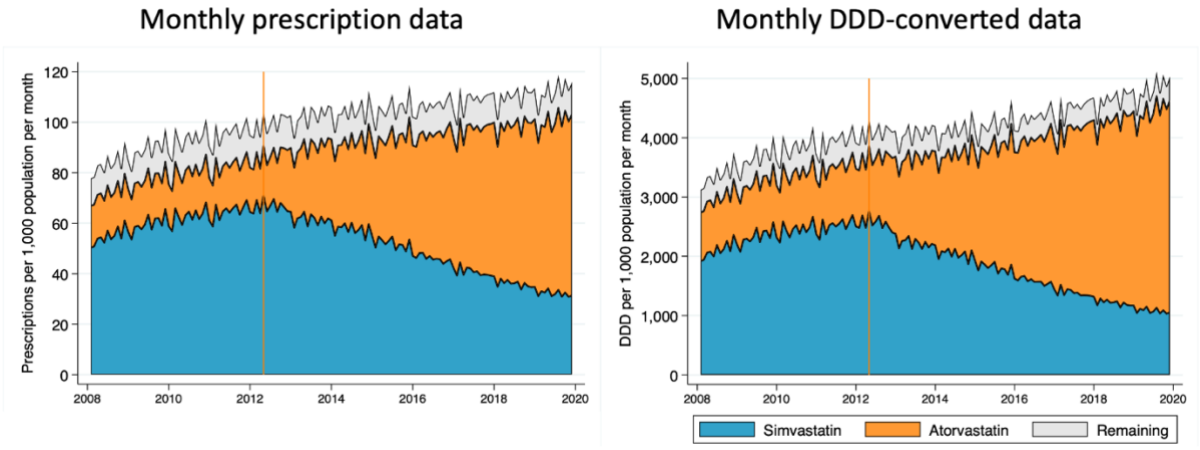
| ATC | Weight | ATC | Weight | ATC | Weight | ATC | Weight | ATC | Weight |
|---|---|---|---|---|---|---|---|---|---|
| C01AA04 | 0.006 | C01EB10 | 0.006 | C03CC01 | 0.006 | C07AA06 | 0.006 | C09AA09 | 0.006 |
| C01AA05 | 0.007 | C01EB15 | 0.006 | C03DA01 | 0.008 | C07AA07 | 0.006 | C09AA10 | 0.006 |
| C01BA01 | 0.006 | C01EB17 | 0.006 | C03DA04 | 0.006 | C07AA12 | 0.006 | C09BA02 | 0.007 |
| C01BA03 | 0.006 | C01EB18 | 0.006 | C03DA05 | 0.006 | C07AB02 | 0.020 | C09BA03 | 0.006 |
| C01BB01 | 0.006 | C02AB01 | 0.006 | C03DB01 | 0.006 | C07AB03 | 0.010 | C09BA04 | 0.006 |
| C01BB02 | 0.006 | C02AC01 | 0.006 | C03DB02 | 0.006 | C07AB04 | 0.006 | C09BA05 | 0.006 |
| C01BC03 | 0.006 | C02AC02 | 0.006 | C03EA01 | 0.008 | C07AB05 | 0.006 | C09BA06 | 0.006 |
| C01BC04 | 0.006 | C02AC05 | 0.006 | C03XA01 | 0.006 | C07AB07 | 0.008 | C09BA08 | 0.006 |
| C01BD01 | 0.006 | C02CA01 | 0.006 | C04AB01 | 0.006 | C07AB08 | 0.006 | C09BB06 | 0.006 |
| C01BD05 | 0.006 | C02CA04 | 0.006 | C04AD03 | 0.006 | C07AB09 | 0.006 | C09BB10 | 0.006 |
| C01BD07 | 0.006 | C02CA06 | 0.006 | C04AE01 | 0.006 | C07AB12 | 0.006 | C09CA01 | 0.007 |
| C01BG07 | 0.006 | C02DB02 | 0.006 | C04AX02 | 0.006 | C07AG01 | 0.006 | C09CA02 | 0.006 |
| C01CA01 | 0.006 | C02DC01 | 0.006 | C04AX21 | 0.006 | C07AG02 | 0.006 | C09CA03 | 0.006 |
| C01CA02 | 0.006 | C02KD01 | 0.006 | C05AA01 | 0.006 | C07FB02 | 0.006 | C09CA04 | 0.006 |
| C01CA03 | 0.006 | C02KX01 | 0.006 | C05AA04 | 0.006 | C08CA01 | 0.009 | C09CA06 | 0.008 |
| C01CA04 | 0.006 | C02KX02 | 0.006 | C05AA08 | 0.006 | C08CA02 | 0.010 | C09CA07 | 0.006 |
| C01CA06 | 0.006 | C02KX03 | 0.006 | C05AE01 | 0.006 | C08CA03 | 0.006 | C09CA08 | 0.006 |
| C01CA07 | 0.006 | C02KX04 | 0.006 | C05AX03 | 0.006 | C08CA05 | 0.006 | C09DA01 | 0.007 |
| C01CA17 | 0.006 | C02KX05 | 0.006 | C05BA01 | 0.006 | C08CA06 | 0.006 | C09DA02 | 0.006 |
| C01CA24 | 0.006 | C03AA01 | 0.008 | C05BA04 | 0.006 | C08CA13 | 0.006 | C09DA03 | 0.006 |
| C01CX08 | 0.006 | C03AA03 | 0.007 | C05BB02 | 0.006 | C08DA01 | 0.006 | C09DA04 | 0.006 |
| C01DA02 | 0.007 | C03AA04 | 0.006 | C05BB04 | 0.006 | C08DB01 | 0.006 | C09DA06 | 0.006 |
| C01DA08 | 0.006 | C03AB01 | 0.006 | C05CA01 | 0.006 | C09AA01 | 0.006 | C09DA07 | 0.006 |
| C01DA14 | 0.009 | C03BA04 | 0.006 | C05CA03 | 0.006 | C09AA02 | 0.013 | C09DA08 | 0.006 |
| C01DX12 | 0.006 | C03BA08 | 0.006 | C05CA04 | 0.006 | C09AA03 | 0.006 | C09DB01 | 0.006 |
| C01DX16 | 0.006 | C03BA11 | 0.006 | C05CA53 | 0.006 | C09AA04 | 0.006 | C09DB02 | 0.006 |
| C01DX22 | 0.006 | C03CA01 | 0.074 | C07AA02 | 0.006 | C09AA05 | 0.008 | C09DX04 | 0.006 |
| C01EA01 | 0.006 | C03CA02 | 0.006 | C07AA03 | 0.006 | C09AA06 | 0.006 | C09XA02 | 0.006 |
| C01EB09 | 0.006 | C03CA04 | 0.006 | C07AA05 | 0.007 | C09AA08 | 0.006 | | |

## Appendix 3.15: Robustness: Replicating the English results using prescription data

Data of the yearly number of prescriptions of each lipid-modifying agent was obtained from various sources. For the years 1998-2003, we obtained it by clicking on "Prescription Cost Analysis: yyyy" where yyyy corresponds to the year in The National Archives (2010). For the years 2004-2018, we clicked on each year and then on "Prescription Cost Analysis yyyy – Tables" to download Excel files at NHS Digital (2023). For 2019, I clicked on "Prescription Cost Analysis 2019 – Statistical summary tables (Excel: 3.8MB)" under NHSBSA (2021b) and for 2020, I clicked on "National summary tables – Calendar year (Excel: 13MB)" under NHSBSA (2021c).

Figure A3.22 displays the English lipid-modifying agents' consumption in two units, using the number of prescriptions on the left and the DDD unit on the right. Both graphs result in a similar picture visually, with just the scale of the units differing.

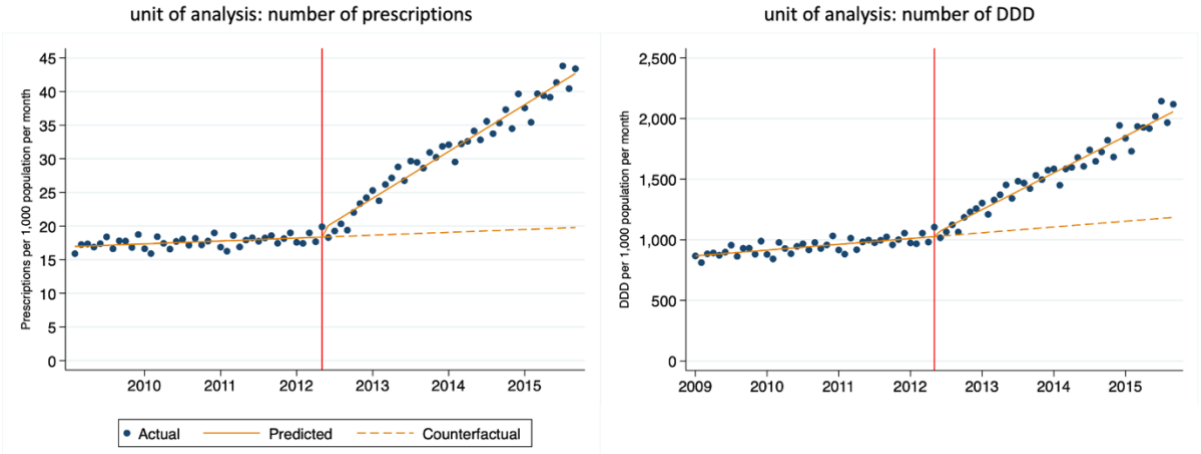
*Appendix figure A3.22: Monthly lipid-modifying agent consumption in England using prescription and DDD units*




Note: Vertical line corresponds to date of atorvastatin's patent expiry.

Applying the interrupted time series also yields similar results as can be seen in appendix figure A3.23 and appendix table A3.26. Both regressions find the overall time trend and the slope change after atorvastatin's patent expiry to be statistically significantly different from zero.

*Appendix figure A3.23: Comparing the impact of patent expiry on English atorvastatin consumption using prescription vis-à-vis DDD units in the Interrupted Time Series analysis*



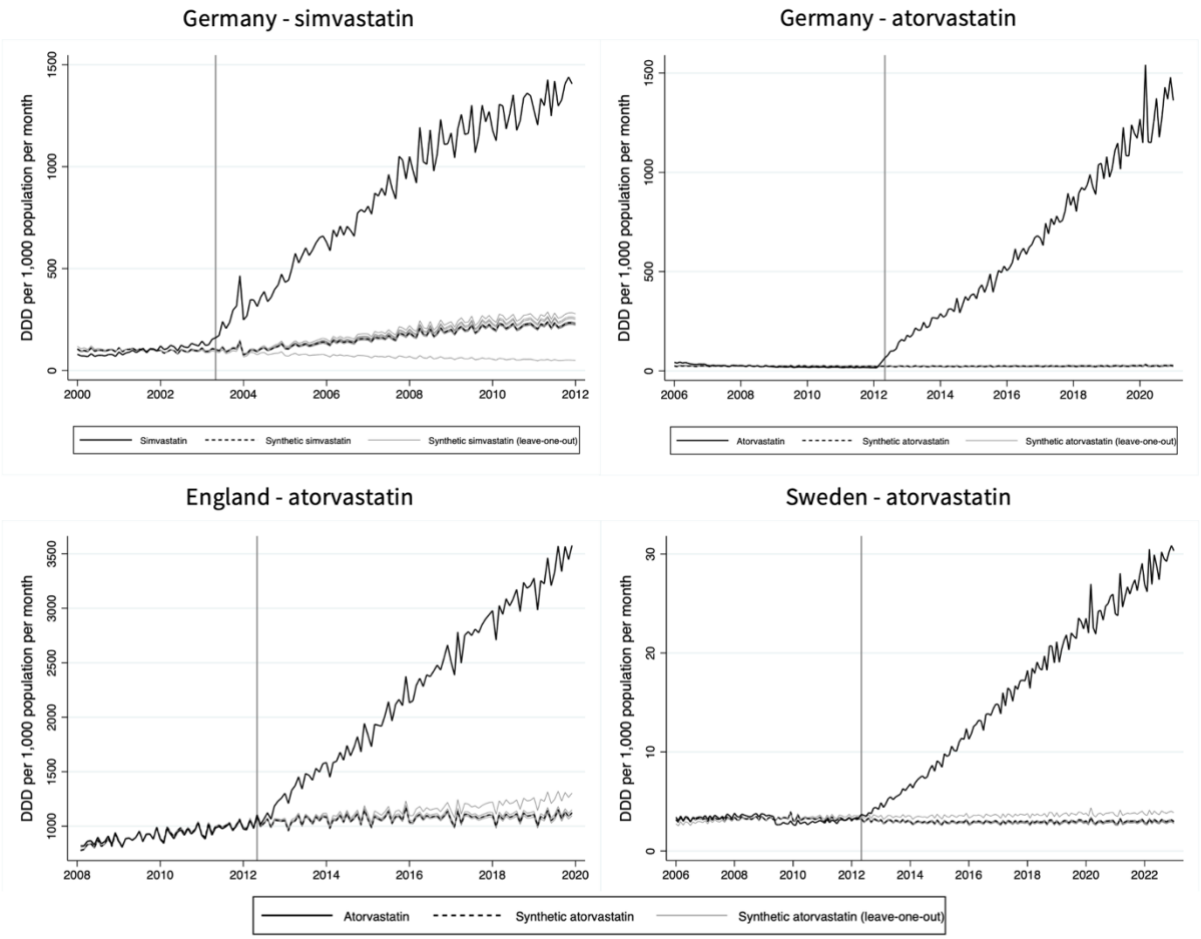

Note: Vertical line corresponds to date of atorvastatin's patent expiry.

*Appendix table A3.27: Comparing the impact of patent expiry on English atorvastatin consumption using prescription vis-à-vis DDD units in the Interrupted Time Series analysis*

| | **DDD (1)** | **Prescriptions (2)** |
|---|---|---|
| time ($\beta_1$) | 3.97*** | 0.036*** |
| | (0.429) | (0.008) |
| expiry ($\beta_2$) | 21.4 | 1.14 |
| | (22.7) | (0.698) |
| time post expiry ($\beta_3$) | 21.3*** | 0.544*** |
| | (0.868) | (0.023) |
| constant ($\beta_0$) | 867*** | 16.9*** |
| | (10.1) | (0.196) |
| Observations | 80 | 80 |

# Appendix 3.16: Robustness: Leave-one-out re-analysis

*Appendix figure A3.24: Leave-one-out re-analysis of the synthetic control method results from Germany, England and Sweden*



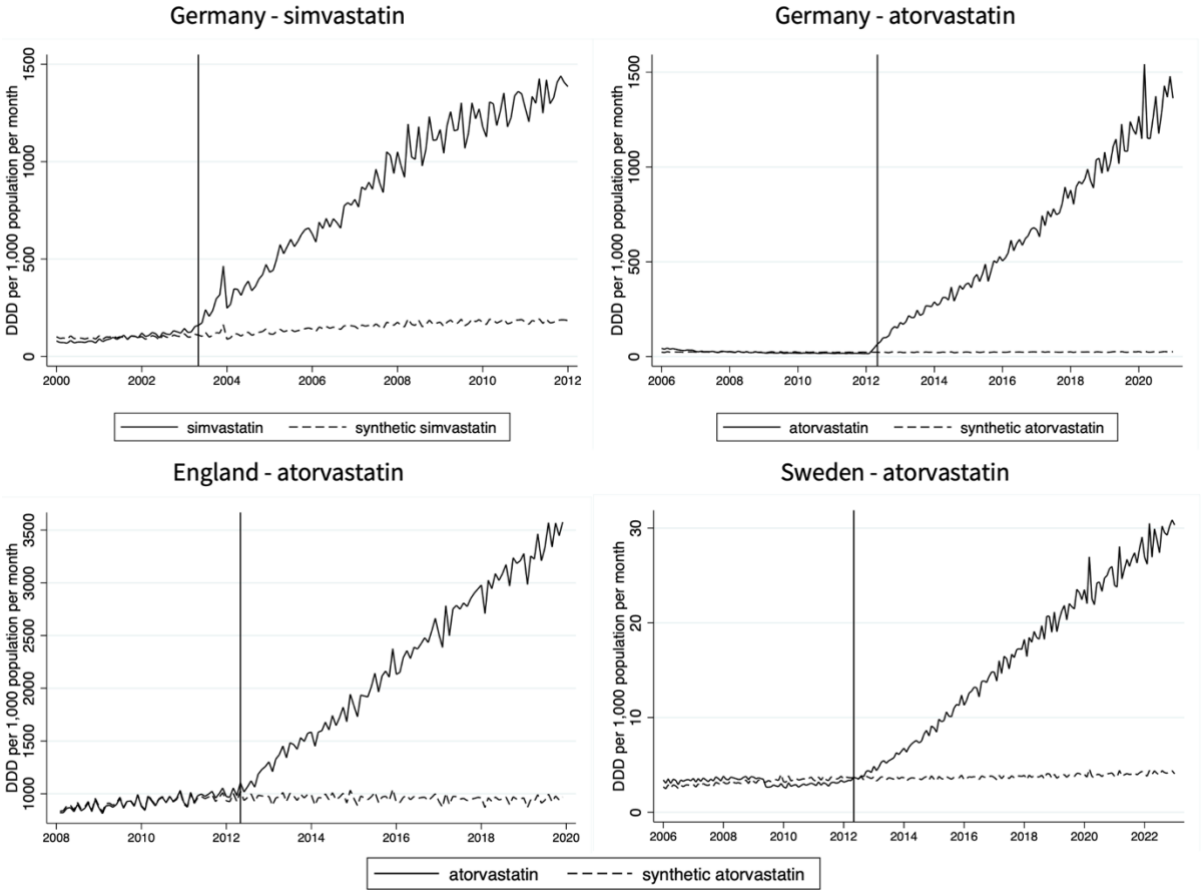

Note: Vertical lines correspond to molecule's patent expiry.

# Appendix 3.17: Robustness: Donor pool of eight molecules

*Appendix table A3.28: Weights of donor pool molecules for the SCM robustness check using only eight molecules*

| Germany simvastatin | | Germany atorvastatin | | England atorvastatin | | Sweden atorvastatin | |
|---|---|---|---|---|---|---|---|
| **ATC** | **Weight** | **ATC** | **Weight** | **ATC** | **Weight** | **ATC** | **Weight** |
| C02AC01 | 0.237 | C02CA01 | 0.744 | C02AC05 | 0.132 | C01DA14 | 0.078 |
| C02CA04 | 0.145 | C02CA04 | 0.012 | C02CA04 | 0.128 | C02CA04 | 0.184 |
| C02CA05 | 0 | C02CA05 | 0 | C03AA01 | 0.123 | C03CA01 | 0.022 |
| C02LA01 | 0.157 | C02CA06 | 0.217 | C03CA01 | 0.126 | C04AE01 | 0.414 |
| C08CA01 | 0.074 | C08CA01 | 0.001 | C08CA01 | 0.121 | C05AA04 | 0.161 |
| C08CA08 | 0.101 | C08CA08 | 0.009 | C08CA02 | 0.129 | C07AB02 | 0.026 |
| C08CA13 | 0.192 | C08CA13 | 0.008 | C09AA03 | 0.126 | C08CA01 | 0.074 |
| C08DA01 | 0.094 | C08DA01 | 0.009 | C09AA05 | 0.113 | C09AA02 | 0.042 |

*Appendix figure A3.25: Consumption of simvastatin and atorvastatin versus their synthetic control using only eight donor pool molecules*




Note: Vertical line correspond to the molecule's patent expiry date.

## Appendix 3.18: Robustness: ITS with reduced time interval (±30 months)

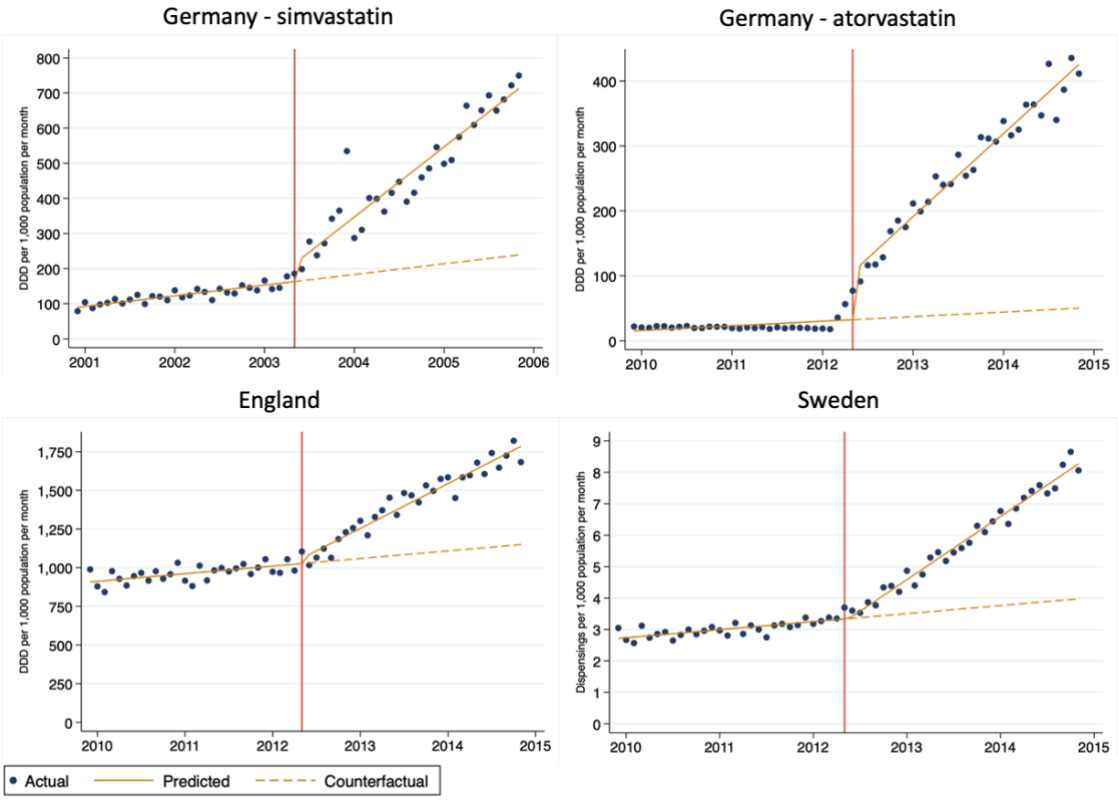
We correct for autocorrelation using Newey-West corrected standard errors and include 3 lags here, too, as $t^{1/4}$ where $t$ corresponds to the number months (t=60) such that $60^{1/4} = 2.783 \approx 3$.

*Appendix table A3.29: Interrupted Time Series results of the impact of patent expiry on drug consumption (± 30 months)*

| | **simvastatin Germany (1)** | **atorvastatin Germany (2)** | **atorvastatin England (3)** | **atorvastatin Sweden (4)** |
|---|---|---|---|---|
| time ($\beta_1$) | 2.53*** | 0.588 | 4.09*** | 0.021*** |
| | (0.232) | (0.468) | (0.646) | (0.004) |
| expiry ($\beta_2$) | 50.9** | 73.3*** | 34.0 | -0.090 |
| | (24.4) | (14.9) | (29.9) | (0.077) |
| time post expiry ($\beta_3$) | 14.1*** | 10.1*** | 20.0*** | 0.146*** |
| | (1.11) | (0.538) | (1.24) | (0.006) |
| constant ($\beta_0$) | 87.4*** | 14.8*** | 904*** | 2.70*** |
| | (2.93) | (5.14) | (11.0) | (0.057) |
| Observations | 60 | 60 | 60 | 60 |

Note: Newey-West standard errors correcting for heteroscedasticity and autocorrelation.

*Appendix figure A3.26: Interrupted time series results of simvastatin and atorvastatin patent expiry (± 30 months)*




Note: The vertical red lines correspond to the date of the molecule's patent expiry.

## 6.4. Appendix references


Agency for Healthcare Research and Quality, 2023. *MEPS – Medical Expenditure Panel Survey.* Available at: meps.ahrq.gov/mepsweb/ (Accessed 4 July 2023).

Aitken, M.L., Berndt, E.R., Bosworth, B., Cockburn, I.M., Frank, R., Kleinrock, M. and Shapiro, B.T., 2013. The regulation of prescription drug competition and market responses: patterns in prices and sales following loss of exclusivity. In *Measuring and Modeling Health Care Costs* (pp. 243-271). University of Chicago Press.

Berndt, E.R., Kyle, M. and Ling, D. 2003. The Long Shadow of Patent Expiration. Generic Entry and Rx-to-OTC Switches. In *Scanner Data and Price Indexes* (pp. 229-273). University of Chicago Press.

Böhm, M., Laufs, U., Hamm, C., Andresen, D., Becker, H.J., Borggrefe, M., Brachmann, J., Dietz, R., et al., 2007. Positionspapier zur Statintherapie*. Clinical Research in Cardiology.* 2, pp. 8-15.

Catapano, A.L., Graham, I., De Backer, G., Wiklund, O., Chapman, M.J., Drexel, H., Hoes, A.W., Jennings, C.S., et al., 2016. 2016 ESC/EAS Guidelines for the Management of Dyslipidaemias. *European Heart Journal*, 37, pp. 2999-3058.

Chapman, S.R., Fitzpatrick, R.W. and Aladul, M.I., 2017. Has cost inhibited the uptake of more potent statins in England?. *Pharmacoepidemiology and Drug Safety*, *26*(8), pp.984-991.

Danish Health Data Authority, 2023. *Medstat.dk*. Available at: medstat.dk/en (Accessed 4 July 2023).

De Backer, G., Ambrosioni, E., Borch-Johnson, K., Brotons, C., Cifkova, R., Dallongeville, J., Ebrahim, S., Faergeman, O., et al., 2003. European guidelines on cardiovascular disease prevention in clinical practice. *European Heart Journal,* 24, pp. 1601-1610.

Department of Health and Social Care, 2000. *National service framework: coronary heart disease*. Available at: www.gov.uk/government/publications/quality-standards-for-coronary-heart-disease-care (Accessed 3 August 2023).

Dietz, R. and Rauch, B., 2003. Leitlinie zur Diagnose und Behandlung der chronischen koronaren Herzerkrankung der Deutschen Gesellschaft für Kardiologie – Herz- und Kreislaufforschung (DGK). *Zeitschrift für Kardiologie,* 92(6), pp. 501-523.

Dißmann, R., Donner-Banzhoff, N., Baum, E., Haasenritter, J., Döhner, W. Werdan, K., Jacobshagen, C., Perings, S., et al., 2019. Nationale Versorgungs-Leitlinie: Chronische KHK. *Ärztliches Zentrum für Qualität in der Medizin.* Available at: www.leitlinien.de/themen/khk/archiv/pdf/khk-5aufl-vers1-lang.pdf (Accessed 3 August 2023).

Donner-Banzhoff, N., Held, K., Laufs, U., Trappe, H.J., Werdan, K. and Zerkowski, H.R. 2006. Nationale Versorgungs-Leitlinie: Chronische KHK. *Ärztliches Zentrum für Qualität in der Medizin.* Available at: www.leitlinien.de/themen/khk/archiv/pdf/khk-lang-1-1.pdf (Accessed 3 August 2023).

Donner-Banzhoff, N., Held, K., Laufs, U., Trappe, H.J., Werdan, K. and Zerkowski, H.R. 2013. Nationale Versorgungs-Leitlinie: Chronische KHK. *Ärztliches Zentrum für Qualität in der Medizin.* Available at: www.leitlinien.de/themen/khk/archiv/pdf/khk-2aufl-vers1-lang.pdf (Accessed 3 August 2023).

Donner-Banzhoff, N., Held, K., Laufs, U., Trappe, H.J., Werdan, K. and Zerkowski, H.R. 2014. Nationale Versorgungs-Leitlinie: Chronische KHK. *Ärztliches Zentrum für Qualität in der Medizin.* Available at: www.leitlinien.de/themen/khk/archiv/pdf/khk-2aufl-vers2-lang.pdf (Accessed 3 August 2023).

Duflos, G. and Lichtenberg, F.R., 2012. Does competition stimulate drug utilization? The impact of changes in market structure on US drug prices, marketing and utilization. *International Review of Law and Economics*, *32*(1), pp.95-109.

Erbel, R., Mudra, H., Sievert, H., Churzidse, S. and Zeller, T. 2012. Kommentar zur Leitlinie (2011) der Europäischen Gesellschaft für Kardiologie zur Diagnose und Therapie der peripheren arteriellen Erkrankungen. *Der Kardiologe,* 6, pp. 302-8.

Estonian State Agency of Medicines, 2013. *Baltic Statistics on Medicines 2010-2012*. Tartu, Estonia: Estonian State Agency of Medicines. Available at: www.zva.gov.lv/en/news-and-publications/publications/baltic-statistics-medicines (Accessed 4 July 2023).

Fiorentini, G., Bruni, M.L. and Mammi, I., 2022. The same old medicine but cheaper: The impact of patent expiry on physicians' prescribing behaviour. *Journal of Economic Behavior & Organization*, *204*, pp.37-68.

Flood, D., Seiglie, J.A., Dunn, M., Tschida, S., Theilmann, M., Marcus, M.E., Brian, G., Norov, B., et al., 2021. The state of diabetes treatment coverage in 55 low-income and middle-income countries: a cross-sectional study of nationally representative, individual-level data in 680 102 adults. *The Lancet Healthy Longevity*, *2*(6), pp.e340-e351.

GKV-Spitzenverband, 2004. *Arzneimittel-Festbeträge: 01.01.2005.* Available at: www.gkv-spitzenverband.de/krankenversicherung/arzneimittel/arzneimittel_festbetraege/festbetragsfestsetzungsbeschluesse_2008_2004/arzneimittel_festbetraege_01_01_2005.jsp (Accessed 27 July 2023).

GKV-Spitzenverband, 2006a. *Arzneimittel-Festbeträge: Anpassung 01.04.2006*. Available at: www.gkv-spitzenverband.de/krankenversicherung/arzneimittel/arzneimittel_festbetraege/festbetragsfestsetzungsbeschluesse_2008_2004/arzneimittel_festbetraege_anpassung_01_04_2006.jsp (Accessed 27 July 2023).

GKV-Spitzenverband, 2006b. *Arzneimittel-Festbeträge – AVWG.* Available at: www.gkv-spitzenverband.de/krankenversicherung/arzneimittel/arzneimittel_festbetraege/festbetragsfestsetzungsbeschluesse_2008_2004/arzneimittel_festbetraege___avwg.jsp (Accessed 27 July 2023).

GKV-Spitzenverband, 2008. *Arzneimittel-Festbeträge: 01.06.2008.* Available at: www.gkv-spitzenverband.de/krankenversicherung/arzneimittel/arzneimittel_festbetraege/festbetragsfestsetzungsbeschluesse_2008_2004/arzneimittel_festbetraege_01_06_2008.jsp (Accessed 27 July 2023)

GKV-Spitzenverband, 2010. *Arzneimittel-Festbeträge: Festbetragsanpassung 01.09.2010.* Available at: www.gkv-spitzenverband.de/krankenversicherung/arzneimittel/arzneimittel_festbetraege/festbetragsfestsetzungsbeschluess_2012_2009/anpassung_09_2010.jsp (Accessed 27 July 2023).

GKV-Spitzenverband, 2012. *Arzneimittel-Festbeträge: Festbetragsanpassung 01.07.2012*. Available at: www.gkv-spitzenverband.de/krankenversicherung/arzneimittel/arzneimittel_festbetraege/festbetragsfestsetzungsbeschluess_2012_2009/arzneimittel_festbetraege_anpassung_07_2012.jsp (Accessed 27 July 2023).

GKV-Spitzenverband, 2014. *Beschluss vom 12.05.2014: Anpassung zum 01.07.2014.* Available at: www.gkv-spitzenverband.de/media/dokumente/krankenversicherung_1/arzneimittel/arzneimittel_festbetraege_1/festbetraege_beschluesse/az_fb_zip_ab_2014/Anpassung_140701.zip (Accessed 27 July 2023).

GKV-Spitzenverband, 2019. *Beschluss vom 29.10.2019: Anwendung der Arzneimittelpreisverordnung ab 01.01.2020*. Available at: www.gkv-spitzenverband.de/media/dokumente/krankenversicherung_1/arzneimittel/arzneimittel_festbetraege_1/festbetraege_beschluesse/az_fb_zip_ab_2014/Umrechnung_20200101.zip (Accessed 27 July 2023).

GKV-Spitzenverband, 2020: *Beschlüsse vom 04.05.2020: Anpassung zum 01.07.2020*. Available at: www.gkv-spitzenverband.de/media/dokumente/krankenversicherung_1/arzneimittel/arzneimittel_festbetraege_1/festbetraege_beschluesse/az_fb_zip_ab_2014/Anpassung_200701.zip (Accessed 27 July 2023).

Gohlke, H., Albus, C., Bönner, G., Darius, H., Eckert, S., Gerber, A., Gohlke-Bärwolf, C., Gysan, D., et al., 2007. Leitlinie Risikoadjustierte Prävention von Herz- und Kreislauferkrankungen. *Deutsche Gesellschaft für Kardiologie.* Available at: leitlinien.dgk.org/files/2007_Leitlinie_Risikoadjustierte_Praevention.pdf (Accessed 3 August 2023).

Gohlke, H., Koenig, W., Schunkert, H., Marx, N., and Hamm, C. 2014. Stellungnahme der Deutschen Gesellschaft für Kardiologie zu den neuen US-Leitlinien zur Verminderung des Atheroskleroserisikos mittels lipidsenkender Therapie. *Der Kardiologe,* 8, pp. 120-124.

Gohlke, H., Kübler, W., Mathes, P., Meinertz, T., Schuler, G., Gysan, D.B., and Sauer, G., 2001. Empfehlungen zur umfassenden Risikoverringerung für Patienten mit koronarer Herzerkrankung, Gefäßerkrankungen und Diabetes. *Zeitschrift für Kardiologie,* 90, pp. 148-149.

Gohlke, H., Kübler, W., Mathes, P., Meinertz, T., Schuler, G., Gysan, D.B., and Sauer, G., 2003. Positionspapier zur Primärprävention kardiovaskulärer Erkrankungen. *Zeitschrift für Kardiologie*, 92(6), pp. 522-524.

Graham, I., Atar, D., Borch-Johnson, K., Boysen, G., Burell, G., Cifkova, R., Dallongeville, J. De Backer, G., et al., 2007. European guidelines on cardiovascular disease prevention in clinical practice: executive summary. European Heart Journal, 28, pp. 2375-2414.

Grundy, S.M., Stone, N.J., Bailey, A. L., Beam, C., Birtcher, K.K., Blumenthal, R.S., Braun, L.T., de Ferranti, S., et al., 2018. 2018 AHA/ACC/AACVPR/AAPA/ ABC/ACPM/ADA/AGS/ APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol. *Circulation*, 139, pp. e1082-e1143.

Harttgen K, Vollmer S. Using an asset index to simulate household income. Economics Letters. 2013 Nov 1;121(2):257-62.

Imai, S., Fushimi, K. and Sundell, K.A., 2018. Impact of new efficacy information on sales of antihypertensive medicines in Japan and Sweden. *Health Policy and Technology*, *7*(2), pp.194-199.

JBS3 Board, 2014. Joint British Societies' consensus recommendations for the prevention of cardiovascular disease (JBS3). *Heart*, 100, pp. ii1-ii67.

Kassenärztliche Bundesvereinigung, 2007. Rahmenvorgaben nach §84 Absatz 7 SGB V und Vereinbarung nach §84 Abs. 7a SGB V – Arzneimittel – für das Jahr 2007. *Deutsches Ärzteblatt,* 104(1-2), pp. A69-72.

Kassenärztliche Bundesvereinigung, 2008a. Rahmenvorgaben nach §84 Absatz 7 SGB V – Arzneimittel – und Hinweise zur Vereinbarung nach §84 Absatz 7a SGB V für das Jahr 2008. *Deutsches Ärzteblatt,* 105(11), pp. A595-596.

Kassenärztliche Bundesvereinigung, 2008b. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2009. *Deutsches Ärzteblatt,* 105(45), pp. A2417-A2419.

Kassenärztliche Bundesvereinigung, 2009. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2010. *Deutsches Ärzteblatt,* 106(45), pp. A2271-A2275.

Kassenärztliche Bundesvereinigung, 2010. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2011. *Deutsches Ärzteblatt,* 107(49), pp. A2465-A2468.

Kassenärztliche Bundesvereinigung, 2011. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2012. *Deutsches Ärzteblatt,* 108(47), pp. A2565-A2569.

Kassenärztliche Bundesvereinigung, 2012. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2013. *Deutsches Ärzteblatt,* 109(48), pp. A2431-A2435.

Kassenärztliche Bundesvereinigung, 2013. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2014. *Deutsches Ärzteblatt,* 110(46), pp. A2232-A2236.

Kassenärztliche Bundesvereinigung, 2014. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2015. *Deutsches Ärzteblatt,* 111(45), pp. A1978-A1982.

Kassenärztliche Bundesvereinigung, 2015. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2016. *Deutsches Ärzteblatt,* 112(45), pp. A1901-A1906.

Kassenärztliche Bundesvereinigung, 2016. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2017. *Deutsches Ärzteblatt,* 113(42), pp. A1885-A1890.

Kassenärztliche Bundesvereinigung, 2017. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2018. *Deutsches Ärzteblatt,* 114(43), pp. A1998-A2003.

Kassenärztliche Bundesvereinigung, 2018. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2019. *Deutsches Ärzteblatt,* 115(42), pp. A1894-A1990.

Kassenärztliche Bundesvereinigung, 2019. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2020. *Deutsches Ärzteblatt,* 116(43), pp. A1977-A1984.

Kassenärztliche Bundesvereinigung, 2020. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2021. *Deutsches Ärzteblatt,* 117(44), pp. A2127-A2134.

Kassenärztliche Bundesvereinigung, 2021. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2022. *Deutsches Ärzteblatt,* 118(42), pp. A1961-A1968.

Kassenärztliche Bundesvereinigung, 2022. Rahmenvorgaben nach §84 Abs. 7 SGB V – Arzneimittel – für das Jahr 2023. *Deutsches Ärzteblatt,* 119(47), pp. A2097-A2104.

Koenig, W., Marx, N., Thiery, J. and Klose, G., 2012. Kommentar zu den neuen Leitlinien (2011) der Europäischen Gesellschaft für Kardiologie zum Management von Dyslipidämien. *Der Kardiologe*. Available at: leitlinien.dgk.org/files/2012_Kommentar_Dyslipidaemie.pdf (Accessed 3 August 2023).

Landmesser, U., Gohlke, H., Hambrecht, R., Kelm, M., Laufs, U., and Marx, N. 2017. Kommentar zu den neuen Leitlinien (2016) der European Society of Cardiology and European Atherosclerosis Society zur Diagnostik und Therapie der Dyslipidämien. *Der Kardiologe,* 11, pp. 295-9.

Lakdawalla, D. and Philipson, T., 2012. Does intellectual property restrict output? An analysis of pharmaceutical markets. *The Journal of Law and Economics*, *55*(1), pp.151-187.

Laufs, U., Donner-Banzhoff, N., Haasenritter, J., Werdan, K., Jacobshagen, C., Fleck, E., Tebbe, U., Hamm, C., et al., 2016. Nationale Versorgungs-Leitlinie: Chronische KHK. *Ärztliches Zentrum für Qualität in der Medizin.* Available at: https://www.leitlinien.de/themen/khk/archiv/pdf/khk-3aufl-vers1-lang.pdf (Accessed 3 August 2023).

Latvian State Agency of Medicines, 2016. *Baltic Statistics on Medicines 2013-2015.* Riga, Latvia: Latvian State Agency of Medicines. Available at: www.zva.gov.lv/en/news-and-publications/publications/baltic-statistics-medicines (Accessed 4 July 2023).

Lithuanian State Agency of Medicines, 2019. *Baltic Statistics on Medicines 2016-2018*. Vilnius, Lithuania: Lithuanian State Agency of Medicines. Available at: www.zva.gov.lv/en/news-and-publications/publications/baltic-statistics-medicines (Accessed 4 July 2023).

Ludt, S., Angelow, A., Baum, E., Chenot, J.F., Donner-Banzhoff, N. and Egidi, G., 2017. *Hausärztliche Risikoberatung zur kardiovaskulären Prävention: S3-Leitlinie*. DEGAM

Marx, N., Anker, S.D., Hammes, H.P. and Tschöpe, C., 2014. Kommentar zu den neuen Leitlinien 2013 der Europäischen Gesellschaft für Kardiologie zu Diabetes, Prädiabetes and kardiovaskulären Erkrankungen. *Der Kardiologe,* 8, pp. 219-22.

Meyer, J., Breithardt, G., Erbel, R. Erdmann, E. Gohlke, H., Hanrath, P. Sonntag, Steinbeck, G., 1998. Leitlinie: Koronare Herzkrankheit / Angina pectoris. *Deutsche Gesellschaft für Kardiologie.* Available at: leitlinien.dgk.org/files/1998_Leitlinie_Koronare_Herzkrankheit_Angina_pectoris.pdf (Accessed 3 August 2023).

Ministry of Health Malaysia, 2006. *Malaysian Statistics on Medicines 2004.* Kuala Lumpur, Malaysia: The National Medicines Use Survey. Available at: www.pharmacy.gov.my/v2/en/documents/malaysian-statistics-medicines.html (Accessed 3 August 2023)

Ministry of Health Malaysia, 2007. *Malaysian Statistics on Medicines 2005.* Kuala Lumpur, Malaysia: The National Medicines Use Survey. Available at: www.pharmacy.gov.my/v2/en/documents/malaysian-statistics-medicines.html (Accessed 3 August 2023)

Ministry of Health Malaysia, 2009. *Malaysian Statistics on Medicines 2006.* Kuala Lumpur, Malaysia: The National Medicines Use Survey. Available at: www.pharmacy.gov.my/v2/en/documents/malaysian-statistics-medicines.html (Accessed 3 August 2023)

Ministry of Health Malaysia, 2010. *Malaysian Statistics on Medicines 2007.* Kuala Lumpur, Malaysia: The National Medicines Use Survey. Available at: www.pharmacy.gov.my/v2/en/documents/malaysian-statistics-medicines.html (Accessed 3 August 2023)

Ministry of Health Malaysia, 2013. *Malaysian Statistics on Medicines 2008.* Kuala Lumpur, Malaysia: The National Medicines Use Survey. Available at: www.pharmacy.gov.my/v2/en/documents/malaysian-statistics-medicines.html (Accessed 3 August 2023)

Ministry of Health Malaysia, 2014. *Malaysian Statistics on Medicines 2009 & 2010.* Kuala Lumpur, Malaysia: The National Medicines Use Survey. Available at: www.pharmacy.gov.my/v2/en/documents/malaysian-statistics-medicines.html (Accessed 3 August 2023)

Ministry of Health Malaysia, 2017. *Malaysian Statistics on Medicines 2011-2014.* Kuala Lumpur, Malaysia: Pharmaceutical Services Division. Available at: www.pharmacy.gov.my/v2/en/documents/malaysian-statistics-medicines.html (Accessed 3 August 2023)

Mach, F., Baigent, C., Catapano, A.L., Koskinas, K.C., Casuala, M., Badimon, L., Chapman, M.J. De Backer, G.G., et al. 2020. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. *European Heart Journal*, 41, pp. 111-88.

Ministry of Health Malaysia, 2020. *Malaysian Statistics on Medicines 2015-2016.* Kuala Lumpur, Malaysia: Pharmaceutical Services Programme. Available at: www.pharmacy.gov.my/v2/en/documents/malaysian-statistics-medicines.html (Accessed 3 August 2023)

Mosca, L., Benjamin, E.J., Berra, K., Bezanson, J.L., Dolor, R.J., Lloyd-Jones, D.M., Newby, K. Piña, I.L., et al. 2011. Effectiveness-Based Guidelines for the Prevention of Cardiovascular Disease in Women – 2011 Update. *Journal of the American College of Cardiology*, 57(12), pp. 1404-23.

National Board of Health and Welfare, 2008. *Nationella riktlinjer för hjärtsjukvård 2008: Beslutsstöd för prioriteringar.* Available at: https://docplayer.se/1764461-Nationella-riktlinjer-for-hjartsjukvard-2008-beslutsstod-for-prioriteringar.html (Accessed 3 August 2023)

National Board of Health and Welfare, 2015. *Hjärtsjukvård Rekommendationer, bedömningar och sammanfattning.* Available at: www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/nationella-riktlinjer/2015-12-5.pdf (Accessed 3 August 2023)

National Board of Health and Welfare, 2018. *Nationella riktlinjer för hjärtsjukvård.* Available at: www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/nationella-riktlinjer/2018-6-28.pdf (Accessed 3 August 2023)

NICE, 2006. *Statins for the prevention of cardiovascular events. Technology Appraisal 94.* London: National Institute for Health and Clinical Excellence.

NICE, 2008a. *Lipid modification. NICE clinical guideline 67.* London: National Institute for Health and Clinical Excellence.

NICE, 2008b. *Cardiovascular disease: identifying and supporting people most at risk of dying early*. London: National Institute for Health and Clinical Excellence. Available at: www.nice.org.uk/guidance/ph15/chapter/1-Recommendations (Accessed 3 August 2023).

NICE, 2014. Cardiovascular disease: risk assessment and reduction, including lipid modification. London: National Institute for Health and Care Excellence. Available at: www.nice.org.uk/guidance/cg181 (Accessed 3 July 2023).

NHS Digital, 2023. *Prescription Cost Analysis*. Available at: digital.nhs.uk/data-and-information/publications/statistical/prescription-cost-analysis (Accessed: 27 June 2023).

NHSBSA, 2021a. Prescription Cost Analysis (PCA) data. Available at: www.nhsbsa.nhs.uk/prescription-data/dispensing-data/prescription-cost-analysis-pca-data (Accessed: 27 June 2023)

NHSBSA, 2021b. *Prescription Cost Analysis – England 2019*. Available at: www.nhsbsa.nhs.uk/statistical-collections/prescription-cost-analysis-england/prescription-cost-analysis-england-2019 (Accessed: 27 June 2023).

NHSBSA, 2021c. *Prescription Cost Analysis – England 2020/21*. Available at: www.nhsbsa.nhs.uk/statistical-collections/prescription-cost-analysis-england/prescription-cost-analysis-england-202021 (Accessed: 27 June 2023).

Norwegian Institute of Public Health, 2021. *Norwegian Prescription Database.* Available at: https://www.norpd.no/Prevalens.aspx (Accessed 2 August 2023).

Perk, J., De Backer, G., Gohlke, H., Graham, I., Reiner, Ž., Verschuren, W.M.M., Albus, C., Benlian, P., et al. 2012. European Guidelines on cardiovascular disease prevention in clinical practice (version 2012). *European Heart Journal,* 33, p. 1635-1701.

Piepoli, M.F., Hoes, A.W., Agewall, S., Albus, C., Brotons, C., Catapano, A.L., Cooney, M.T., Corrà, U., et al. 2016. 2016 European Guidelines on cardiovascular disease prevention in clinical practice. *European Heart Journal*, 37, pp. 2315-81.

Public Health Scotland, 2020. *Prescription Cost Analysis.* Available at: www.isdscotland.org/Health-Topics/Prescribing-and-Medicines/Community-Dispensing/Prescription-Cost-Analysis/ (Accessed: 2 August 2023)

Reinecke, H., Brandenburg, V., Dominiak, P., Flöge, J., Galle, H., Geiger, B., Grabensee, F., de Haan, K., et al. 2006. Empfehlungen zur Diagnostik und Behandlung von Patienten mit koronarer Herzkrankheit und Niereninsuffizienz. *Clinical Research in Cardiology,* 1, pp. 103-117.

Reiner, Ž., Catapano, A., De Backer, G., Graham, I., Taskinen, M.R., Wiklund O., Agewall, S., Alegria, E., et al. 2011. ESC/EAS Guidelines for the management of dyslipidaemias. *European Heart Journal,* 32, pp. 1769-1818.

Riley L, Guthold R, Cowan M, et al. The World Health Organization STEPwise approach to noncommunicable disease risk-factor surveillance: methods, challenges, and opportunities. *American Journal of Public Health* 2016; **106**: 74–78.

Rydén, L., Grant, P.J., Anker, S.D., Berne, C., Cosentino, F., Danchin, N., Deaton, C., Escaned, J., et al., 2013. ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD. *European Heart Journal,* 34, pp. 3035-87.

Savin, N.E. and White, K.J., 1977. The Durbin-Watson test for serial correlation with extreme sample sizes or many regressors. *Econometrica: Journal of the Econometric Society*, pp.1989-1996.

Scheller, B., Levenson, B., Joner, M., Zahn, R, Klauss, V., Naber, C. Schächinger, V. and Elsässer, A., 2005. Supplementband Primärprävention kardiovaskulärer Erkrankungen*. Zeitschrift für Kardiologie,* 94(Suppl 3). Available at: herzmedizin.de/fuer-aerzte-und-fachpersonal/leitlinien/leitlinien-archiv/supplementband_primaerpraevention_kardiovaskulaerer_erkrankungen.html (Accessed 3 August 2023).

Schneider, B., Donner-Banzhoff, N., Kühlein, T., Baum, E., Haasenritter, J., Egidi, G., Rubin, D., Werdan, K., et al. 2022. Nationale Versorgungs-Leitlinie: Chronische KHK. *Ärztliches Zentrum für Qualität in der Medizin.* Available at: www.leitlinien.de/themen/khk (Accessed 3 August 2023).

Schuler, G.C., Koenig, W., Adams, V. and Gohlke, H., 2013. Kommentar zu den neuen Leitlinien (2012) der Europäischen Gesellschaft für Kardiologie zur kardiovaskulären Prävention. *Der Kardiologe*, 7, pp. 251-60.

Smith, S.C., Allen, J., Blair, S.N., Bonow, R.O., Brass, L. M., Fonarow, G.C., Grundy, S.M., Hiratzka, L., et al., 2006. AHA/ACC Guidelines for Secondary Prevention for Patients With Coronary and Other Atherosclerotic Vascular Disease: 2006 Update. *Circulation*, 113, pp. 2363-2372.

Smith, S.C., Benjamin, E.J., Bonow, R.O., Braun, L.T., Creager, M.A., Franklin, B.A. Gibbons, R.J., Grundy, S.C., et al. 2011. AHA/ACCF Secondary Prevention and Risk Reduction Therapy for Patients with Coronary and Other Atherosclerotic Vascular Disease: 2011 Update. *Journal of the American College of Cardiology*, 58(23), pp. 2432-46.

Smith, S.C., Blair, S.N., Bonow, R.O., Brass, L. M., Cerqueira, M. D., Dracup, K., Fuster, V., Gotto, A. et al., 2001. AHA/ACC Guidelines for Preventing Heart Attack and Death in Patients With Atherosclerotic Cardiovascular Disease: 2001 Update. *Circulation*, 104, pp. 1577-1579.

STATA. STATA Survey Data Reference Manual Release 18 [Internet]. Texas: StataCorp LLC; 2023. Available from https://www.stata.com/manuals/svy.pdf (accessed May 5, 2023).

Stone, N.J., Robinson, J.G., Lichtenstein, A.H., Bairey Merz, N., Blum, C.B., Eckel, R.H., Goldberg, A.C., Gordon, D., et al., 2013. 2013 ACC/AHA Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults. *Circulation*, 129(suppl 2): pp. S1-S45.

Tendera, M., Aboyans, V., Bartelink, M.L., Baumgartner, I., Clément, D., Colleg, J.P., Cremonesi, A., De Carlo, M., et al. 2011. ESC Guidelines on the diagnosis and treatment of peripheral artery diseases. *European Heart Journal*, 32, pp. 2851-2906.

The National Archives, 2010. *Health care statistics: Pharmacies and prescriptions*. Available at: webarchive.nationalarchives.gov.uk/ukgwa/20120503222906/http://www.dh.gov.uk/en/Publicationsandstatistics/Statistics/StatisticalWorkAreas/Statisticalhealthcare/DH_4086488 (Accessed: 27 June 2023).

TLV, 2023a. Sök priser och beslut i databasen – atorvastatin. Available at: www.tlv.se/beslut/sok-priser-och-beslut-i-databasen.html?product=atorvastatin&tab=3 (Accessed 3 August 2023).

TLV, 2023b. Sök priser och beslut i databasen – simvastatin. Available at: www.tlv.se/beslut/sok-priser-och-beslut-i-databasen.html?product=simvastatin&tab=3 (Accessed 3 August 2023).

WHO. WHO Package of Essential Noncommunicable (PEN) Disease Interventions for Primary Health Care [Internet]. Geneva: World Health Organization; 2020. Available from apps.who.int/iris/rest/bitstreams/1301957/retrieve (accessed May 5, 2023).

Wood D., De Backer, G., Faergeman, O., Graham, I., Mancia, G., Pyörälä, K., 1998a. Prevention of coronary heart disease in clinical practice: Recommendations of the Second Joint Task Force of European and other Societies on Coronary Prevention. *Atherosclerosis*, 140: pp. 199-270.

Wood, D., Durrington, P., McInnes, G., Poulter, N., Rees, A., Wray, R., 1998b. Joint British recommendations on prevention of coronary heart disease in clinical practice. *Heart*, 80(Suppl. 2), pp. S1–29.

Wood, D., Wray, R., Paulter, N., Williams, B., Kirby, M., Patel, V., Durrington, P., Reckless, J., et al., 2005. JBS 2: Joint British Societies' guidelines on prevention of cardiovascular disease in clinical practice. *Heart*, 91(Suppl V): pp. v1-v52.

# 7. Author contribution statement

For all three essays, my contributions correspond to that of a first author (when listing authors non-alphabetically).

*Essay 1: The impact of grants in combination with school-based management trainings on primary education: a cluster-randomized trial in Northern Nigeria*

Essay 1 was published jointly with Kehinde Elijah Owolabi, Folake Olatunji-David, Niyi Okunlola and Sebastian Vollmer in the Journal of Development Effectiveness (Ochmann, Owolabi, Olatunji-David, Okunolola and Vollmer, 2022). My colleagues Lisa Bogler and Ann-Charline Weber conceptualized the study together with Sebastian Vollmer. The data collection of the baseline data was conducted by Lisa Bogler, Ann-Charline Weber, Kehinde Elijah Owolabi, Niyi Okunlola and Sebastian Vollmer. The endline data collection was conducted by Kehinde Elijah Owolabi, Niyi Okunlola, Sebastian Vollmer and myself. I analyzed the data and wrote the first draft with substantial contributions from Sebastian Vollmer.

*Essay 2: Diagnostic testing for hypertension, diabetes and hypercholesterolaemia in low-income and middle-income countries: a cross-sectional study of data from 994 185 individuals from 57 nationally representative surveys*

Essay 2 was published jointly with Isabelle von Polenz, Maja-Emilia Marcus, Michaela Theilmann, David Flood, Kokou Agoudavi, Krishna Kumar Aryal, Silver Bahendeka, Brice Bicaba, Pascal Bovet, Luisa Campos Caldeira Brant, Deborah Carvalho Malta, Albertino Damasceno, Farshad Farzadfar, Gladwell Gathecha, Ali Ghanbari, Mongal Gurung, David Guwatudde, Corine Houehanou, Dismand Houinato, Nahla Hwalla, Jutta Adelin Jorgensen, Khem B Karki, Nuno Lunet, Joao Martins, Mary Mayige, Sahar Saeedi Moghaddam, Omar Mwalim, Kibachio Joseph Mwangi, Bolormaa Norov, Sarah Quesnel-Crooks, Negar Rezaei, Abla M. Sibai, Lela Sturua, Lindiwe Tsabedze, Roy Wong-McClure, Justine Davies, Pascal Geldsetzer, Till Bärnighausen, Rifat Atun, Jennifer Manne-Goehler and Sebastian Vollmer. I cite the author contribution statement as stated on the published manuscript with the Lancet Global Health, where the initials of co-authors were used as abbreviations. “SO, IP, RA, JM-G, and SV conceived the study. M-EM, MT, DF, JD, PG, TB, RA, JM-G and SV led the data curation, with

input from KA, KKA, SB, BB, PB, LCCB, DCM, AD, FF, GG, AG, MG, DG, CH, DH, NH, JAJ, KBK, NL, JM, MM, SSM, OM, KJM, BN, SQ-C, NR, AMS, LS, LT and RW-M. SO led the formal analysis with substantial contributions from IP, RA, JM-G, and SV. SO and IvP wrote the first draft of the Article, with substantial contributions from RA, JM-G, and SV. All authors had full access to the data, which were verified by SO, IVP, RA, JM-G and SV. All authors provided input on several versions of the article, read and approved the final version, and had final responsibility for the decision to submit for publication." (Ochmann et al., 2023)

*Essay 3: The impact of patent expiry on statin consumption: A synthetic control analysis*

Essay 3 is joint work with Gabriele Gradl, Martin Schulz and Sebastian Vollmer. Sebastian Volmer and I co-conceived the study. I led the data curation with contributions from Gabriele Gradl and Martin Schulz. I conducted the analyses and wrote the first draft with substantial contributions from Gabriele Gradl, Martin Schulz and Sebastian Vollmer.