The effect of payment method and multimorbidity on health and healthcare utilisation

Helen Hayes, Jonathan Stokes, Søren Rud Kristensen, Matt Sutton

Abstract

Purpose

Three types of payment methods have been introduced across European countries in attempts to encourage better, more integrated care of persons with multimorbidity: pay-for-performance; pay-for-coordination; and an all-inclusive payment method. We examine whether there are differences in the way these payment methods affect health and healthcare use in persons with multimorbidity.

Design/Methodology/Approach

Using individual-level survey data from twenty European countries, we examine unadjusted differences in average outcomes for the years 2011-2015 by whether countries adopted new payment methods for integrated care. We then test for a differential effect for multimorbid persons using linear, individual random effects regressions, including country and time fixed effects and clustering standard errors at the country level.

Findings

We find little effect of varying payment methods on key outcomes for multimorbid individuals despite the theoretical predictions and the rhetoric in many policy documents.

Originality

This paper is the first to study the effects of payments for integration on the dimensions and populations these schemes intend to affect; health and healthcare use at the individual level for multimorbid individuals.

Research limitations/implications

Policymakers should bear in mind that the success of the payment method relies on the specific design of the incentives and their implementation. New effective models of care and how to incentivise these for multimorbid patients is an ongoing research priority.

Keywords: Multimorbidity, integrated care, pay-for-performance, pay-for-coordination, bundled payment, global payment

Introduction

Multimorbidity, the coexistence of two or more chronic conditions, is a growing concern for most health systems, and for patients with chronic conditions the presence of multiple as opposed to one single condition is now considered the norm (World Health Organization, 2016). The presence of multimorbidity in a patient has adverse health effects, with multimorbid patients experiencing decreased quality of life, increased mortality rates, and a higher risk of developing further chronic mental health conditions (Fortin *et al.*, 2007; Mujica-Mota *et al.*, 2015; Rizzuto *et al.*, 2017). Multimorbidity also increases healthcare utilisation; increases hospital admissions and time spent in hospital (Fortin *et al.*, 2007; Smith *et al.*, 2012).

Fragmented care can be costly. This is true in terms of the inefficient use of resources without proper coordination between providers with risk of duplication; and in its effect on patients, where more difficult or complicated access to the care they need may have negative effects on future health. The concept of integrated care involves a move away from a single disease management approach to more coordinated, person-centred, holistic care to cope with changing demographics and increasing numbers of people with multimorbidity. Integrated care is seen as a high priority for policymakers in countries such as the UK, featuring prominently in the NHS long term plan (NHS England, 2019), and globally, as set out by the World Health Organization (World Health Organisation, 2015).

Fragmented care is likely to pose a particular risk to those with multimorbidity who frequently see multiple care providers across multiple sectors. A single disease management approach to care may not be appropriate in treating patients with complex comorbidities and payment schemes which incentivise this approach could emphasise inappropriate aspects of care for multimorbid patients. This could have adverse effects on the quality of care, health, and healthcare utilisation (Boyd *et al.*, 2005; Sinnott and Bradley, 2015). Multimorbid patients are therefore those most susceptible to the negative effects of fragmented care, and the potential benefits of integrated care.

Current payment schemes have been described as barriers to coordination of care and to incentivising providers, payers and patients to increase their efforts to do so (Struckmann *et al.*, 2017). A variety of new payment methods have therefore been introduced across European countries in attempts to encourage better, more integrated care of persons with multimorbidity (Busse and Mays, 2008).

Tsiachristas et al. (2016) describe three classifications of payment mechanisms to incentivise integration: pay-for-coordination, which involves payment for different types of providers to co-ordinate care for specific services; pay-for-performance, which provides a financial incentive for care providers to meet certain performance measures and goals in the treatment and outcomes of care for chronic conditions; and, the all-inclusive payment scheme which comprises of (i) bundled payments, where a single payment is given for a specific condition which could involve multiple services and providers; and (ii) global payments, which cover a specified group of patients' healthcare costs (not only disease-specific), over a fixed period of time (Tsiachristas *et al.*, 2016).

The intuition for each of these payment types is that if coordinating care becomes a direct consideration for a profit maximising provider, payer or patient's economic success, then they will be incentivised to dedicate resources towards integrating care. In doing so, this should theoretically improve patient healthcare outcomes, and either reduce healthcare utilisation to control costs, or increase preventative utilisation (and so, health) with improved access, depending on the incentives of the scheme.

Previous research has used national-level data from 25 countries to estimate the effects of integrated payment methods on trends in health care expenditure for the entire population. The authors found that pay-for-coordination and the all-inclusive methods decreased outpatient expenditure growth, while pay-for-performance decreased overall hospital and administrative expenditure growth (Tsiachristas *et al.*, 2016). More general comparisons of payment reform, Moreno-Serra and Wagstaff (2010) look at the shift from historical budgets to fee-for-service (FFS) or patient-based payments (PBP), and find that both payment types increased national health spending, however had mixed effects on inpatient admissions and average length of stay, and limited effects on mortality. Wubulihasimu, Brouwer and van Baal (2016), also look at the adoption of FFS and PBP schemes and find that FFS increases the growth rate of healthcare output, and PBP has positive effects on life expectancy at age 65. The current literature does not examine effects of integrated care payment mechanisms on the health and use of healthcare services by patients with multimorbidity.

In this paper, we use individual-level survey data to evaluate differences in health and healthcare utilisation in countries which adopt these various payment methods, compared to countries where no new payment method was introduced. As outlined above, those patients with multimorbidity are the primary targets of integrated care, those most vulnerable to the negative effects of fragmented systems and where we would expect to identify any benefits inferred from the changing payment methods. Therefore, we evaluate these payment methods by exploiting the gap between individuals with and without multimorbidity.

Methods

Data

We obtained microdata for twenty countries using the Survey of Health, Ageing and Retirement in Europe (SHARE) waves 1 (Börsch-Supan, 2019a), 2 (Börsch-Supan, 2019b), 3 (SHARELIFE) (Börsch-Supan, 2019c), 4 (Börsch-Supan, 2019d), 5 (Börsch-Supan, 2019e) and 6 (Börsch-Supan, 2019f) for 2004 to 2015. SHARE collects responses from the native-speaking population aged 50 and over who permanently reside in each country. Spouses and partners are also interviewed independent of age if they reside in the same household. The data collection involves computer-assisted face-to-face interviews supplemented by a self-completion questionnaire. Our analysis contains information on 111,446 individuals, 6,624 of whom were members of all six waves.

The UK does not participate in SHARE, but has a similar survey with overlapping questions and design, the English Longitudinal Study of Ageing (ELSA) (Banks *et al.*, 2019). We therefore supplemented the dataset with data from the first six waves of ELSA, which contains data for 2002 to 2013. ELSA draws sample members over the age of 50 from respondents to the Health Survey for England (HSE). ELSA also involves a computer-assisted personal interview and is followed up by a self-completion questionnaire which takes place every two years of the study. At waves 3, 4 and 6 the study was refreshed with new participants from HSE to ensure a representative sample of people in their 50s and to maintain sample size. In our analysis we use information on 8,463 individuals, 3,059 of whom were members of all six waves.

The final sample contains information on 311,510 observations, provided by 119,909 individuals with full information on the control variables. 9,683 of these individuals were members of all six waves. However, over half of the total respondents (64,889) only appeared in 1 or 2 waves. 32,661 respondents appeared in only one wave.

Payment approaches

Table I depicts how we applied the payment methods classification to the available respondent-level data in each year/wave in each country. Some countries have not introduced payment methods aimed specifically at encouraging integrated care at any stage in the study period, acting as controls. As in Tsiachristas et al. (2016), we assume the payment scheme adopted was maintained in years where there is no information available towards the end of the sample. Grey shaded cells indicate missing data in SHARE/ELSA in that country and year.

Table I: Payment plans in place by year and country, based on Tsiachristas et al. (2016), p425.

Country	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Wave			**1**	**1**	**1/2**	**2**	**3**	**3**	**3/4**	**3/4**	**4**	**5**		**6**
England			PFP	PFP	PFP	PFP	PFP	PFP	PFP	PFP	PFP	PFP	PFP	PFP
Austria				PFC	PFC	PFC	PFC	PFC	PFC	PFC	PFC	PFC	PFC	PFC
Germany		PFC/ALL	PFC/ALL	PFC/ALL	PFC/ALL	PFC/ALL	PFC/ALL	PFC/ALL	PFC/ALL	PFC/ALL	PFC/ALL	PFC/ALL	PFC/ALL	PFC/ALL
Sweden														
Netherlands										ALL	ALL	ALL	ALL	ALL
Spain														
Italy														
France				PFC	PFC	PFC	PFC	PFP/PFC	PFP/PFC	PFP/PFC	PFP/PFC	PFP/PFC	PFP/PFC	PFP/PFC
Denmark						PFC	PFC	PFC	PFC	PFC	PFC	PFC	PFC	PFC
Greece														
Switzerland														
Belgium														
Czech Republic														
Poland														
Ireland														
Luxembourg														
Hungary	PFC	PFC	PFC	PFC	PFC	PFC	PFC	PFP	PFP	PFP	PFP	PFP	PFP	PFP
Portugal										PFP	PFP	PFP	PFP	PFP
Slovenia														
Estonia					PFP	PFP	PFP	PFP	PFP	PFP	PFP	PFP	PFP	PFP

Notes: Cells shaded grey indicate that no data are available from SHARE/ELSA in that country/year. Empty cells indicate that no dedicated payment schemes were in place.

Pay-for-performance

The pay-for-performance schemes included in this study target primary care providers; either GPs or multidisciplinary teams. They provide a financial incentive to meet integrated care related performance targets.

Pay for coordination

The pay-for-coordination schemes promote provider coordination activity targeting at the primary care, regional and insurer level.

All-inclusive

For the all-inclusive payment method, the bundled payment is mainly aimed at primary care providers while the risk adjusted global payment made patients with multimorbidity and integrated care more attractive to insurers.

Expected impact of payment methods on patient outcomes

We hypothesise each of the three types of payment incentives for integration to be associated with greater coordination between care providers leading to greater integration. In turn, we expect this greater integration of care to be associated with better patient outcomes (self-assessed health) and reduced healthcare utilisation, particularly secondary care utilisation as integrated care models aim to a more preventative healthcare model. With a lack of integrated care measures, we are not able to observe the actual change in provider behaviour which we hypothesise to be the mechanism between payments and our outcomes. However, Table II exhibits a more detailed explanation of the specific payment structures in place for each of the payment method-adopting countries as outlined in Tsiachristas et al. (2016). We expect that multimorbidity patients will be those most affected by the integration of care and are therefore interested in the differential effect for multimorbid patients of being in one of the payment method adopting countries, compared to a country with no payment method in place. We do not have ex ante expectations about the relative strength of the changes in outcomes induced by different types of payment.

Table II Setup of payment method by country, based on Tsiachristas et al. (2016) p.423.

	Reform	Target of reform	Integrated care incentive
Pay-for-performance (only)			
England (PFP)	Quality and Outcomes Framework 2004	GPs	• Integrated care related performance indicators included in PFP contracts • GPs rewarded with a possible 25-30% salary increase • Public reporting of GP practice QOF performance
Hungary (PFP)	Performance based payment 2009	GPs	• Bonus payment system for GPs based on performance and quality indicators • €1.1 million initially invested for GP bonuses increased to €10.9 million by 2012
Portugal (PFP)	Performance compensations 2006	Multi-disciplinary primary care teams	• Performance compensations relating to preventative care and chronic conditions targeted at vulnerable and high risk patients • Multidisciplinary teams entitled rather than independent GPs
Estonia (PFP)	Performance based payment 2009	GPs	• PFP system for ambulatory care • Targeted GP performance in disease prevention, monitoring chronic conditions and increased professional competency • GPs could get up to €255 per month on top of usual payment
Pay-for-coordination (without all-inclusive)			
Austria (PFC)	Health reform act 2005	Social health insurers and regional governments	• State level financial pools of 1-2% of combined budget of social health insurers with that of regional governments • Pooled funds created for ICPs (mainly DMPs)
		GPs	• GPs received an initial premium (€53) to enrol a patient in DMP and quarterly payment (€25) to supplement the FFS
Denmark (PFC)	Administrative reform 2007	Region Municipalities	• 14 Danish regions replaced by 5, and responsibilities reallocated • Coordination between regions and municipalities: regions provided with €70 million to implement ICPs and municipalities responsible for co-financing • 15% of regional healthcare budgets for ICPs • Municipalities co-financing hospital care
France (PFC and PFP)	Health Insurance Act 2004	*(PFC element)* GPs	• GPs receive a supplement of €40 per patient enrolled in a DMP • System of gate keeping for primary care • Support for GP training and education
	CAPI 2009	*(PFP element)* GPs	• GPs could get €6000 annually (30% of base salary) for adequately registered patient records and following evidence based guidelines
All-inclusive (any)			
Netherlands (ALL)	Bundled payment	Mainly primary care providers	• Bundled payment for diabetes DMPs (2007) and COPD and cardiovascular disease (2010) • Chronic care coordinated by care groups, implementing DMPs • Insurers negotiate a predefined fee to cover all care of a patient with a particular chronic condition for a year • Performance is a factor in price negotiations
Germany (ALL and PFC)	Risk Structure compensation Act 2002	*(ALL element)* Social Health insurers	• Risk-adjusted global payment • More attractive to insure patients with chronic conditions and set up DMPs (DMP registration became an additional risk adjuster) • Insurers could benefit up to €1000 per patient by implementing DMPs • Updated 2009 taking individual morbidity indicators into account so insurers receive higher remuneration for chronically ill patients, regardless of DMP enrolment
		(PFC element) Social Health insurers	• Insurers setting up DMPs for chronically ill patient receive additional €153 for coordination costs
		GPs	• GPs provided with €75 per patient per year for coordination and coordination costs GPs had incentive to be contracted by insurers to increase their package of provided services reimbursed by insurers

Notes: DMPs: Disease Management Programs; GPs: General Practitioners; ICPs: Integrated Care Program; PFP: Pay-for-performance; PFC: Pay-for-coordination; ALL: All-inclusive. .

Outcome measures

Health outcomes

We focus on three health variables: self-assessed health, limitations with activities of daily living, and quality of life (CASP). We briefly outline how these variables are coded and the data collection process.

1. *Self-assessed health*
   For self-assessed health, respondents are asked to rate their health on a five-point scale. For the purpose of our analysis this is coded giving the range 1 [Poor] – 5 [Excellent].
2. *Limitations with activities of daily living*
   The limitations with activities of daily living index describes the number of limitations respondents report to have over six everyday self-care activities ranging from 0 [no limitations] to 6 [most limitations]. These limitations include problems with dressing, walking, grooming, eating, getting in or out of bed and using the toilet. For the purpose of analysis this scale is reversed such that a higher score indicates fewer limitations, and better health, for consistency with the other health outcomes.
3. *Quality of life (CASP)*
   The CASP scale used in SHARE (Mehrbrodt, Gruber and Wagner, 2019) comprises four sub-scales: control, autonomy, self-realisation and pleasure. Respondents are presented with 12 items as questions or statements and assessed on a four-point Likert scale. The overall score is the sum of these 12 items, giving a range of scores between 12 and 48 with high scores corresponding to greater wellbeing.

Utilisation outcomes

For our three utilisation measures we use: the total number of times an individual has seen or spoken to any doctor in the last 12 months; the total number of these doctor contacts with a GP; and GP visits as a proportion of overall doctor visits.

1. *Total number of times seen or spoken to a medical doctor in the last 12 months*
   For the number of times seen or spoken to a doctor in the last 12 months, respondents are asked about how many times in total they have seen or talked to a medical doctor about their health (range 0-98+) and are asked to exclude dentist visits and hospital stays, but include emergency room or outpatient clinic visits.
2. *Total number of doctor contacts (from 1, above) with a GP in the last 12 months*
   For the number of total contacts with a doctor (above), respondents are asked how many of these contacts were with a general practitioner or with a doctor at their health care centre (range 0-98+).
3. *GP contacts as a proportion of overall doctor contacts*
   Respondents were asked how many of these total contacts with a doctor (above) were with a general practitioner or with a doctor at their health care centre (range 0-98+). We use this to create a variable indicating the number of contacts with a GP as a proportion of overall contacts with a medical doctor (range 0-1).

Sample sizes vary across outcomes as data for some indicators are not collected in certain countries or some waves of SHARE or ELSA. Additionally, the response rate is lower across countries for contacts with a doctor, particularly for GP contacts as this is a sub-question only if answering positively to having had any doctor contacts.

Analysis

Table I is referenced throughout this section to illustrate the problem of missing data and how this influences the methods employed in our analysis.

Descriptive analysis

We first present descriptive statistics for all outcome and control variables for the full sample. Next, we look at differences in average outcomes by payment type for the years 2011-2015 only, as fairly complete information is available in these years for all countries (see Table I). As there are two countries which have multiple payment mechanisms in place in these years (Germany: pay-for-coordination and all-inclusive; France: pay-for-performance and pay-for-coordination), we assign countries to four payment groups. These are: no specific payments for integration; pay-for performance only; pay-for-coordination without all-inclusive; and any country which implements the all-inclusive payment method. Table II provides further details of the countries in each payment method grouping.

Regression analysis

Next, we more formally test these differences with regressions using all available time points. Our choice of analytic design is affected by the lack of observations prior to the introduction of the payment methods in some cases (as shown in Table I). Ideally, we would use the variation in payment methods across time and countries in a traditional policy evaluation, for example using difference-in-differences (Goodman-Bacon, 2018). One possibility would be to restrict our analysis to the countries for which sufficient pre-reform data exist (only two out of the nine payment method adopting countries have more than one year of pre-reform data). However, in that case our analysis would not provide generalisable findings about the impact of payment methods across countries.

Instead, we look to test whether there is a differential multimorbidity effect under any of the three payment methods with equation (1). In essence, we draw on the assumption from the literature that multimorbid patients will be those who primarily benefit from any incentivised integration of care. Therefore, we allow non-multimorbid patients to act as controls within each country and focus on exploring the gaps between these two groups across countries and over time. This allows our analysis to include all available data points. In summary, our identification strategy focuses on multimorbid patients, and the variation in outcomes across payment method for these patients.

$$Y_{ict} = \beta_1 MM_{it} + \beta_2 MM * PFP_{it} + \beta_3 MM * PFC_{it} + \beta_4 MM * ALL_{it} + \beta_5 PFP_{ct} + \beta_6 PFC_{ct} + \beta_7 ALL_{ct} + \beta_8 X_{it} + \gamma_c + \tau_t + \alpha_i + \varepsilon_{it} \quad (1)$$

Y_{ict} corresponds to the six health and utilisation outcomes (each estimated separately) for individual i in country c at time t. MM_{it} represents a dummy variable equal to 1 if an individual is multimorbid in time t, and 0 otherwise (see Appendix I for a full list of the chronic conditions included in the definition of the multimorbidity variable). $MM * PFP_{it}$ refers to a dummy variable equal to 1 if an individual is multimorbid and there is a pay-for-performance scheme in place in time t, and 0 otherwise. $MM * PFC_{it}$ and $MM * ALL_{it}$ are similarly constructed dummy variables for pay-for-coordination and the all-inclusive payment method. β_2, β_3 and β_4 capture the differential multimorbidity effect under each payment method, relative to the multimorbidity effect in countries that did not adopt one of these payment methods. PFP_{it}, PFC_{it} and ALL_{it} are dummy variables equal to 1 if an individual is in a country with a pay-for-performance, pay-for-coordination or all-inclusive payment scheme in place, respectively, in time t, and 0 otherwise. We include the three main effects for the payment methods to ensure that the interaction effects measure only the differential

multimorbidity effect of interest. X_{it} refers to a set of socioeconomic controls: gender (using a dummy variable for male); age categories (20-59 years [base category]; 60-69 years; 70-79 years; 80-89 years; and 90 years and over); and level of education (less than high school [base category], high school graduate, some college and above), we also interact gender with age. All regressions control for country fixed effects (γ_c) estimated by C-1 country dummies, time fixed effects (τ_t) estimated by T-1 interview year dummies and include an unobserved individual random effect (α_i).

Out of the 119,909 individuals observed in our data, only 9,683 individuals are followed in all waves, so using a balanced panel would exclude the majority of the sample. A number of countries are only included for one wave, and a number of the countries adopting an innovative payment method do not change payment method during the observed period. Given this, we are unable to include fixed effects at the respondent level. Instead, we estimate equation (1) using linear, individual random effects regressions for all outcomes to allow for cases of repeated observations of individual respondents.

Robustness checks

We rerun the analysis on self-assessed health using an ordered logistic regression, given the Likert-scale nature of this measure.

In Hungary, previous to the pay-for-performance scheme was a pay-for-coordination scheme involving a coordinator who was incentivised financially to manage coordination of care across primary, secondary and tertiary care which was aborted in 2008 (with pay-for-performance scheme introduced the following year). In our analysis we only observe data for 2011; the potential negative effect of the removal of the pay-for-coordination scheme may be captured in the effect of introducing pay-for-performance in Hungary. As a robustness check, we repeat the analysis excluding Hungary.

We assume that the classification of payments by Tsiachristas et al. (2016) persists once implemented. As a robustness check, we estimate the models without the years where we have assumed this continuation.

Results

Descriptive analysis

Table III: Descriptive statistics for dependent and independent variables for the full sample

	Mean	Min	Max	Count
Dependent variables				
Health outcomes				
Self-assessed health	2.895 (1.096)	1	5	282902
Quality of life (CASP)	37.364 (6.322)	12	48	220218
Number of limitations with activities of daily living (reversed scale)	5.748 (0.868)	0	6	276210
Utilisation outcomes				
How often seen or talked to medical doctor last 12 months	6.700 (9.681)	0	98	179430
How many of these contacts with general practitioner	4.251 (6.978)	0	98	124123
Proportion of doctor visits with a GP in the last 12 months	0.733 (0.321)	0	1	102382
Independent variables				
Multimorbidity	0.530	0	1	311510
Age	66.386	21	111	311510
Male	0.447	0	1	311510
Education		1	4	311510
Less than high school	0.40			
High school	0.34			
Some college	0.06			
College and above	0.21			

Notes: Descriptive statistics are provided for all observations across all countries and time points (N*T), with full information on the controls. Standard deviations are included in parentheses underneath the dependent variables. All health outcomes are coded such that higher is "better". Observations vary by outcome as some indicators are not collected in certain countries or some waves of SHARE/ELSA.

Error! Reference source not found. presents descriptive statistics on key outcome and control variables for the full sample. The proportion of individuals in our sample with multimorbidity is high, but unsurprising given the average age of our sample (Barnett *et al.*, 2012).

Multimorbidity effect

Table IV presents average outcomes for the years 2011-2015. This shows that, as expected, compared to healthy individuals, multimorbid patients have worse self-assessed health, quality of life, and limitations with activities of daily living. Focusing on individuals in countries with no specific payment method to incentivise integration, multimorbid patients have more overall medical doctor contacts and GP contacts, and a lower proportion of their doctor contacts with a GP (i.e. relatively more of their contacts are with other medical doctors) compared with non-multimorbid individuals.

Table IV: Comparing health and utilisation measures by payment classification and multimorbidity for the years 2011-2015 only

	No payment method			PFP only			PFC without ALL			ALL		
	MM	Non-MM	Difference	MM	Non-MM	Difference	MM	Non-MM	Difference	MM	Non-MM	Difference
Health outcomes												
Self-assessed health	3.325	2.395	-0.929***	3.032	2.062	-0.970***	3.548	2.576	-0.972***	3.249	2.379	-0.869***
	(0.957)	(0.949)		(1.028)	(0.893)		(0.989)	(1.014)		(0.961)	(0.877)	
Quality of life (CASP)	38.826	34.875	-3.951***	37.591	33.161	-4.430***	41.250	37.747	-3.502***	41.060	37.899	-3.161***
	(5.510)	(6.497)		(5.536)	(6.265)		(4.746)	(6.180)		(4.727)	(5.857)	
Limitations with activities of daily living	5.919	5.552	-0.368***	5.919	5.494	-0.425***	5.939	5.610	-0.329***	5.922	5.636	-0.286***
	(0.499)	(1.174)		(0.453)	(1.161)		(0.407)	(1.061)		(0.483)	(1.017)	
Utilisation outcomes												
How often seen or talked to a medical doctor last 12 months	4.152	9.380	5.228***	3.626	7.758	4.133***	4.335	8.878	4.544***	5.136	10.706	5.570***
	(6.687)	(11.659)		(6.709)	(9.966)		(6.289)	(10.559)		(8.700)	(13.348)	
How many of these contacts with a GP	2.126	6.081	3.954***	2.434	5.165	2.731***	2.452	5.555	3.103***	1.692	4.657	2.966***
	(3.891)	(9.182)		(5.886)	(7.372)		(3.676)	(7.045)		(4.538)	(7.220)	
Proportion of doctor visits with a GP in the last 12 months	0.746	0.721	-0.025***	0.764	0.745	-0.019***	0.736	0.739	0.003	0.660	0.602	-0.057***
	(0.332)	(0.305)		(0.341)	(0.302)		(0.322)	(0.293)		(0.360)	(0.330)	

Notes: * p<0.1; ** p<0.05; *** p<0.01. All health outcomes are coded such that higher is "better" health. MM=multimorbid. Non-MM=non-multimorbid. Average values are for the years 2011-2015 only as these provide fairly complete payment information.

Payment method differences

Differences between the multimorbid and non-multimorbid group for health outcomes are consistent across all payment methods (individuals with multimorbidity have worse health on average than those without). Considering volume of doctor contacts, multimorbid individuals also have more contacts with any medical doctor overall, and more contacts with a GP. These differences in means are all statistically significant. No payment method, pay-for-performance only, and the all-inclusive method have a lower proportion of doctor contacts with a GP. In pay-for-coordination countries, excluding those with all-inclusive payments as well, the average multimorbid individual has a higher proportion of doctor visits with a GP compared to non-multimorbid individual, though this difference is not statistically significant at the 10% level.

Relative to average health outcomes in countries adopting no dedicated payment method, pay-for-performance only countries have generally poorer health outcomes, while pay-for-coordination countries have better health outcomes. These differences between outcomes under payment methods apply to both multimorbid and non-multimorbid groups.

The all-inclusive payment method has both the highest average total doctor visits and the lowest of these visits with a GP, on average, for both multimorbid and non-multimorbid groups.

Regression analysis

Table V shows the same multimorbidity effect as the unadjusted result: worse health outcomes, more healthcare utilisation, and a proportionally higher emphasis on other medical doctor usage compared to GPs. The magnitude of the differences is congruent with the unadjusted differences.

The quality of life (CASP) score is reduced by 0.649 ($p<0.05$) and the limitations with activities of daily living score is reduced by 0.076 ($p<0.01$) for multimorbid individuals in pay-for-performance countries. These effects are representative of close to 10 and 9 percent of a standard deviation (6.322 and 0.868), respectively.

There is an additional increase in the number of overall doctor contacts for multimorbid patients in all-inclusive countries of 0.931 ($p<0.01$). This relative increase in doctor contacts signifies close to 10 percent of a standard deviation of overall doctor contacts (9.682).

There is an increase in the proportion of contacts with a GP for pay-for-performance countries for multimorbid patients of 0.02 (6 percent of a standard deviation), though this result is only significant at the 10% level and is not supported in either of the robustness analyses.

Table V: Individual random effects regressions results showing the effects of multimorbidity and differential multimorbidity effect under the three payment methods, on health and utilisation

	Health outcomes			Utilisation outcomes		
	Self-assessed health	Quality of life (CASP)	Limitations with activities of daily living	How often seen or talked to medical doctor last 12 months	How many of these contacts with a GP	Proportion of doctor visits with a GP in the last 12 months
MM	-0.756***	-3.169***	-0.239***	4.427***	2.644***	-0.038***
	(0.012)	(0.176)	(0.018)	(0.336)	(0.306)	(0.008)
MM*PFP	-0.016	-0.649**	-0.076***	-0.461	-0.337	0.001
	(0.045)	(0.311)	(0.024)	(0.582)	(0.588)	(0.012)
MM*PFC	-0.039	0.239	0.014	0.047	0.066	0.020*
	(0.038)	(0.309)	(0.036)	(0.506)	(0.396)	(0.012)
MM*ALL	0.007	0.213	0.004	0.931**	-0.039	-0.023
	(0.027)	(0.302)	(0.029)	(0.444)	(0.635)	(0.025)
Observations	282902	220218	276210	179430	124123	102382

Notes: Robust, country-clustered standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01

All models control for a pay-for-performance, pay-for-coordination, and all-inclusive payment main effect dummy, as well as education, gender, age categories, gender interacted with age categories, and country and year fixed effects. All health outcomes are coded such that higher is "better" health. MM=multimorbidity, PFP=pay-for-performance, PFC=pay-for-coordination, ALL=all-inclusive payment.

1.1 Robustness checks

Use of an ordered logistic regression to estimate the effect on self-assessed health supports the main finding of no differential multimorbidity effect across payment types (see Appendix II).

Results for the analysis without Hungary can be found in Appendix III. Results support the main analysis in terms of sign, significance and magnitude of the effects for the main multimorbidity effect and the group effect for pay-for-coordination. For pay-for-performance and pay-for coordination, the direction of the effect across all outcomes is unchanged. The coefficient on the proportion of doctor contacts with a GP is no longer significant at the 10% level for pay-for-coordination countries. The significant effect of worse limitations with activities of daily living within pay-for-performance countries is robust to this sensitivity analysis. The coefficient on quality of life is no longer significant at the 10% level. There is now a significant decrease in the number of doctor contacts in the last 12 months (-0.908, $p<0.05$) and a decrease in the number of contacts with a GP (0.820, $p<0.05$).

Results for the analysis excluding years where payment information was assumed to be consistent can be found in Appendix IV. Results for the main multimorbidity effect support those in the main analysis in terms of sign and significance. The slight changes in the magnitude of the differences are reasonable given the loss of data points. Three of the significant results by payment group in the main analysis are supported: a lower quality of life and limitation with activities of daily living score for multimorbid individuals in pay-for-performance countries, and more overall doctor contacts for multimorbid individuals in all-inclusive countries. The pay-for-coordination result on the proportion of doctor contacts with a GP is no longer statistically significant.

Discussion

Principal findings and potential mechanisms

There is little evidence of any major differences for multimorbid individuals in payment method adopting countries. The three main statistically significant effects we find are a worse quality of life and a worse activities of daily living score for multimorbid individuals in pay-for-performance countries (-0.649, $p<0.05$; -0.076, $p<0.01$), and increased numbers of overall doctor contacts for multimorbid patients in all-inclusive countries (0.931, $p<0.01$). These results are economically meaningful but are not in the expected direction. The quality of life result is not supported in analysis excluding Hungary. The lack of effect found here could be supported by motivation crowding theory, which suggests that offering monetary incentives may crowd out intrinsic motivation. Many of the pay-for-performance and pay-for-coordination schemes studied here involve bonuses offered to physicians to meet integrated care related targets or enrol patients in disease management programs (DMPs). It could be that for an altruistic physician who feels intrinsically motivated to help their patient, offering a monetary incentive to perform a task may crowd out their innate enjoyment of it and undermine their performance (Burgess and Ratto, 2003; Frey, 2017).

Pay-for-performance is associated with worse quality of life and a worse activities of daily living score for multimorbid individuals. In many cases, the integrated care related pay-for-performance targets were part of wider pay-for-performance schemes, where not all targets relate to co-ordinating care. Having multiple, potentially conflicting, pay for performance targets to meet may result in adverse health effects (Boyd *et al.*, 2005; Weiss *et al.*, 2014)

The all-inclusive payment is associated with an increase in overall doctor contacts for multimorbid individuals. This finding is consistent with previous studies which found an overall increase in total health expenditure after introduction of the Dutch bundled payments (Struijs *et al.*, 2012). As above, these incentivised a number of single DMPs rather than aimed at multimorbidity more generally. Incentivising a number of disease-specific DMPs simultaneously may also incentivise increased duplication and double-payment, particularly for those with more than one incentivised condition. Additionally, if patient experience (which we were not able to measure) was improved by better coordinated care, patients may be more likely to use it; or in the short term, with improved organisation, patients may be more quickly referred onto the required service – the Dutch bundled payments only included primary care activity, for instance, so there may have been an adverse incentive for primary care practitioners to refer patients to other setting reducing their own costs and able to keep more of the bundled payment. Although we did not see a significant change in proportion with a GP, there was a non-significant decreasing trend.

Although we do not find evidence for significant improvements to health and utilisation outcomes by payment type in the countries studied here, it may be that the issue is in the implementation rather than the theory behind the payment types themselves. For example, some incentive schemes may be more effective than others due to the size of the reward, as a significant percentage of income has to be variable before providers can be expected to change their behaviour (Busse *et al.*, 2010)

Strengths and weaknesses of the study

To our knowledge, we are the first to study the effects of payments for integration on the dimensions and populations these schemes intend to affect; health and healthcare use at the individual level for multimorbid individuals. The multimorbid population is likely most sensitive to the effects of the

incentives we study. The prevalence of multimorbidity in populations is increasing (Koné Pefoyo *et al.*, 2015) and will continue to increase as populations age. Furthermore, multimorbidity is not just an issue prevalent in the elderly; in a recent English study 27.2% of patients (aged 18 and over) had more than one chronic condition (Cassell *et al.*, 2018). More research is needed on how payment methods can be used to improve the care of multimorbid patients (The Academy of Medical Sciences, 2018). This study provides evidence which aims to fill that gap.

The challenge of any comparative countries study is finding the balance of describing institutional detail in sufficient detail while allowing findings to be generalizable across countries. In our analysis we describe the specific setup of the payment method in each country to understand the mechanisms through which they should impact on our outcome variable. Within the payment classifications we outline there is variation in how the types of payments are applied. Different ways of applying the specific payments, incentivising specific types of activity for specific groups, is likely to vary the results within each category. However, in this paper we are interested in overarching themes of the effects of payment methods on integrated care for multimorbid patients. We find little overall effect.

The results may not represent causal effects of introducing these payment methods for the multimorbid population because data availability means that we cannot compare outcomes before and after the introduction of new payment methods. Moreover, health inequities are driven primarily by pre-determined, social determinants of health, and healthcare has only limited effects (Marmot and Wilkinson, 2005). However, the associations between payment methods and outcomes are based on individual-level data from large, multi-country longitudinal surveys. We are unable to include individual fixed effects which would control for many of these pre-determined characteristics, although we do include individual random effects.

Assuming a linear scale in our analysis could be problematic in the case of self-assessed health, as this ignores the ordered categorical nature of this measure and assumes the distance between each category is equal. However, we have checked the sensitivity of this result to using an ordered logistic regression and the results are unchanged: we find no differential multimorbidity effect under any of the three payment methods on self-assessed health.

The cross-sectional survey weights provided by SHARE and ELSA are not applied to this analysis. We use individual random effects to account for the same individuals in multiple (but not necessarily all) waves, and longitudinal weights are only available for the very small number of people contributing to all six surveys. Furthermore, the methodology of the weights for ELSA changed between waves 1-2 and waves 3-6 (NatCen Social Research, 2018), and England is one of the key intervention countries. For SHARE, these cross-sectional weights are based on age, gender and regional areas (SHARE, 2020), so in any case would not help with our estimation strategy as the weights do not consider multimorbidity status. Instead, we incorporate age and gender and other, observable covariates in the regression models. If sampling probabilities vary exogenously instead of endogenously, use of weighting may be unnecessary for consistency and harmful for precision (Solon, Haider and Wooldridge, 2015). Thus, the additional benefit of weighting a sample based on observables which are included as covariates is unclear.

Attrition will introduce bias into our estimates if the reason for missing data is related to our outcome measures. For example, if patients drop out of the sample because of illness and do not follow up in consecutive waves this may bias our results when looking at health outcomes. Banks Muriel and Smith (2011) find that attrition does not appear related to prior disease prevalence in ELSA. Further, analysis

from a survey administered to refusing respondents from the German Wave 4 showed little evidence for health status related nonresponse bias (Börsch-Supan *et al.*, 2013).

This study uses self-reported survey data. The three health measures we focus on are therefore likely to be correlated, but are different ways of measuring health since health is multi-dimensional (World Health Organization, 2020). Using this survey data might have some issues for recording utilisation. In using utilisation data which relies on an individual's recollection of doctor contacts over the period of a year, there is a risk of measurement error as patients may not recall correctly. However, we would expect the same level of bias in countries where payments changed and where they did not. Furthermore, using patient-reported utilisation indicator gives a measure of service use across all potential services where an individual may contact a medical doctor, although the most expensive part of a health system, in-patient admissions, are not included in these contacts as respondents are not asked to consider hospital stays.

Results in relation to other studies

Tsiachristas et al. (2016) finds that the payment methods analysed in this study can have significant effects on healthcare expenditure. In our study we find no real evidence for improvements to health and healthcare utilisation.

Payment is only one aspect of care integration. As in Tsiachristas (2016) we were unable to control for non-financial policies but assumed that they were equally likely to occur in treatment and control countries. This is a strong assumption, and in reality, there is likely to be heterogeneity in the use of non-financial policies for integration. It is possible that other methods of integrating care could be driving the differences across payment methods. However, including country fixed effects will control for non-financial policies of integrating care which are in place over the whole period, as well as general attitude towards integration policies. There are mixed results of the effect of integrated care more generally on utilisation (Nolte and Pitchforth, 2014). Case management, a popular method of integrating care for multimorbid patients is found to have a positive but small effect on self-assessed health in the short term only and no significant effect on utilisation of primary or secondary care (Stokes *et al.*, 2015). So, it is not clear whether the finance schemes have little effect on stimulating integrated care activity, or whether the activity it stimulates only generates a weak signal in turn.

The literature shows that incentives work best when rewards are directed at individual physicians or small groups (Eijkenaar *et al.*, 2013) which may be why we see more significant associations for schemes like pay-for-performance which is largely targeted at GPs. This is supported by Kristensen et al. (2016) who find that pay-for-performance payments are more effective at the department level rather than the hospital level.

The single disease approach is an important problem in ageing populations with multimorbidity quickly becoming the norm rather than the exception in care of chronic patients (Violan *et al.*, 2014). Overlap of multiple disease management programs for patients with multimorbidity can complicate care (Juul-Larsen *et al.*, 2017). Following evidence-based guidelines for patients with multiple chronic conditions, which schemes such as pay-for-performance encourage adherence to, may result in adverse drug and care interactions (Boyd *et al.*, 2005; Marengoni and Onder, 2015). The integrated care models which have been most successful to date are longer-term and disease management models for single conditions (Rocks *et al.*, 2020), which might in fact have negative effects for multimorbid patients. Another relevant area of the literature to consider is multitasking (Holmstrom and Milgrom, 1991) when some conditions are not included in the schemes. If chronic conditions are

substitutes in terms of doctors' effort, incentivising providers to dedicate time and resources to care for a certain chronic condition may result in less focus on other chronic conditions.

Mechanisms and implications for clinicians or policymakers

When interpreting the results of this analysis, policymakers should bear in mind that the success of any payment method relies on the specific design of incentives and their implementation. These include the responsiveness to the scheme, the size of the incentives in place, who is targeted (provider, payer or patient), and how the aspects of care which are incentivised interact with those which are not (i.e. does incentivising one area come at the cost of disincentivising another to the overall detriment of care). Stokes et al. (2018) provide a typology of payments for integrated care to encourage improved financial incentives.

Policymakers should also consider the reason for using payment mechanisms to implement integrated care programmes. Though there is evidence to suggest that these payment methods impact costs (Tsiachristas *et al.*, 2016), if the focus is on improving health and controlling utilisation, more careful consideration of the setup of these payment methods may be required. Overall, payment methods are not a panacea to the outcomes sought from integrated care.

Conclusions

This paper is the first to look at the three classifications of payment methods for integrated care outlined by Tsiachristas et al. (2016) with respect to health and utilisation outcomes. It considers the impact of payment methods to stimulate integrated care on the multimorbid population who stand to benefit most from their implementation. There is little difference in the multimorbidity effect by payment method despite the rhetoric in the theoretical literature. The effectiveness of models of care and how to incentivise these for multimorbid patients is an ongoing research priority (Stokes *et al.*, 2017; MacMahon, 2018).

Reference list

Banks, J. *et al.* (2019) 'English Longitudinal Study of Ageing: Waves 0-8, 1998-2017'. UK Data Service. SN: 5050, p. 29th Edition. doi: 10.5255/UKDA-SN-5050-16.

Banks, J., Muriel, A. and Smith, J. P. (2011) 'Attrition and health in ageing studies: evidence from ELSA and HRS', *Longitudinal and Life Course Studies*. Bristol University Press, 2(2). doi: 10.14301/llcs.v2i2.115.

Barnett, K. *et al.* (2012) 'Epidemiology of multimorbidity and implications for health care, research, and medical education: A cross-sectional study', *The Lancet*. Lancet Publishing Group, 380(9836), pp. 37–43. doi: 10.1016/S0140-6736(12)60240-2.

Börsch-Supan, A. *et al.* (2013) 'Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE)', *International Journal of Epidemiology*. Oxford Academic, 42(4), pp. 992–1001. doi: 10.1093/ije/dyt088.

Börsch-Supan, A. (2019a) 'Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 1'. Release version: 7.0.0. SHARE-ERIC. doi: 10.6103/SHARE.w1.700.

Börsch-Supan, A. (2019b) 'Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 2'. Release version: 7.0.0. SHARE-ERIC. doi: 10.6103/SHARE.w2.700.

Börsch-Supan, A. (2019c) 'Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 3 - SHARELIFE'. Release version: 7.0.0. SHARE-ERIC. doi: 10.6103/SHARE.w3.700.

Börsch-Supan, A. (2019d) 'Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 4'. Release version: 7.0.0. SHARE-ERIC. doi: 10.6103/SHARE.w4.700.

Börsch-Supan, A. (2019e) 'Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5'. Release version: 7.0.0. SHARE-ERIC. doi: 10.6103/SHARE.w5.700.

Börsch-Supan, A. (2019f) 'Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6'. Release version: 7.0.0. SHARE-ERIC. doi: 10.6103/SHARE.w6.700.

Boyd, C. M. *et al.* (2005) 'Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: Implications for pay for performance', *Journal of the American Medical Association*. American Medical Association, pp. 716–724. doi: 10.1001/jama.294.6.716.

Burgess, S. and Ratto, M. (2003) 'The Role of Incentives in the Public Sector: Issues and Evidence', *Oxford Review of Economic Policy*, 19(2), pp. 285–300. doi: 10.1093/oxrep/19.2.285.

Busse, R. *et al.* (2010) *Tackling chronic disease in Europe. Strategies, interventions and challenges*, *Political Science*. Available at: http://www.euro.who.int/en/publications/abstracts/tackling-chronic-disease-in-europe-strategies,-interventions-and-challenges-2010 (Accessed: 7 May 2020).

Busse, R. and Mays, N. (2008) 'Paying for Chronic Disease Care', in Nolte, E. and McKee, M. (eds) *Caring for People with Chronic Conditions: A Health System Perspective*. Open University Press, pp. 195–221.

Cassell, A. *et al.* (2018) 'The epidemiology of multimorbidity in primary care: A retrospective cohort study', *British Journal of General Practice*. Royal College of General Practitioners, 68(669), pp. e245–e251. doi: 10.3399/bjgp18X695465.

Eijkenaar, F. *et al.* (2013) 'Effects of pay for performance in health care: A systematic review of systematic reviews', *Health Policy*, pp. 115–130. doi: 10.1016/j.healthpol.2013.01.008.

Fortin, M. *et al.* (2007) 'Multimorbidity and quality of life: A closer look', *Health and Quality of Life Outcomes*, 5. doi: 10.1186/1477-7525-5-52.

Frey, B. (2017) 'Policy consequences of pay-for-performance and crowding-out', *Journal of Behavioral Economics for Policy*, 1(1), pp. 55–59.

Goodman-Bacon, A. (2018) 'Difference-in-Differences with Variation in Treatment Timing', *National Bureau of Economic Research*. doi: 10.3386/w25018.

Holmstrom, B. and Milgrom, P. (1991) 'Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design', *Journal of Law, Economics & Organization*, 7.

Juul-Larsen, H. G. *et al.* (2017) 'Prevalence and overlap of Disease Management Program diseases in older hospitalized patients', *European Journal of Ageing*. Springer Verlag, 14(3), pp. 283–293. doi: 10.1007/s10433-017-0412-9.

Koné Pefoyo, A. J. *et al.* (2015) 'The increasing burden and complexity of multimorbidity disease epidemiology - Chronic', *BMC Public Health*. BioMed Central Ltd., 15(1), pp. 1–11. doi: 10.1186/s12889-015-1733-2.

Kristensen, S. R., Bech, M. and Lauridsen, J. T. (2016) 'Who to pay for performance? The choice of organisational level for hospital performance incentives', *European Journal of Health Economics*. Springer Verlag, 17(4), pp. 435–442. doi: 10.1007/s10198-015-0690-0.

MacMahon, S. (2018) *Multimorbidity: a priority for global health research*. London, UK.

Marengoni, A. and Onder, G. (2015) 'Guidelines, polypharmacy, and drug-drug interactions in patients with multimorbidity: A cascade of failure', *BMJ (Online)*. BMJ Publishing Group. doi: 10.1136/bmj.h1059.

Marmot, M. and Wilkinson, R. (2005) *Social determinants of health*. Available at: www.ihepa.in (Accessed: 28 September 2020).

Mehrbrodt, T., Gruber, S. and Wagner, M. (2019) *SHARE: Scales and Multi-Item Indicators*.

Moreno-Serra, R. and Wagstaff, A. (2010) 'System-wide impacts of hospital payment reforms: Evidence from Central and Eastern Europe and Central Asia', *Journal of Health Economics*, 29(4), pp. 585–602. doi: 10.1016/j.jhealeco.2010.05.007.

Mujica-Mota, R. E. *et al.* (2015) 'Common patterns of morbidity and multi-morbidity and their impact on health-related quality of life: evidence from a national survey', *Quality of Life Research*. Kluwer Academic Publishers, 24(4), pp. 909–918. doi: 10.1007/s11136-014-0820-7.

NatCen Social Research (2018) *ELSA User Guide to the Main Interview Datasets - Waves 1 to 8*.

NHS England (2019) 'The NHS Long Term Plan'. Available at: www.longtermplan.nhs.uk (Accessed: 5 May 2020).

Nolte, E. and Pitchforth, E. (2014) 'What is the evidence on the economic impacts of integrated care?' World Health Organization.

Rizzuto, D. *et al.* (2017) 'Effect of Chronic Diseases and Multimorbidity on Survival and Functioning in Elderly Adults', *Journal of the American Geriatrics Society*. Blackwell Publishing Inc., 65(5), pp. 1056–1060. doi: 10.1111/jgs.14868.

Rocks, S. *et al.* (2020) 'Cost and effects of integrated care: a systematic literature review and meta-analysis', *European Journal of Health Economics*. Springer, 21(8), pp. 1211–1221. doi: 10.1007/s10198-020-01217-5.

SHARE (2020) *SHARE Release Guide 7.1.0*.

Sinnott, C. and Bradley, C. P. (2015) 'Multimorbidity or polypharmacy: two sides of the same coin?', *Journal of comorbidity*. SAGE Publications, 5(1), pp. 29–31. doi: 10.15256/joc.2015.5.51.

Smith, S. M. *et al.* (2012) 'Managing patients with multimorbidity: Systematic review of interventions in primary care and community settings', *BMJ (Online)*. doi: 10.1136/bmj.e5205.

Solon, G., Haider, S. J. and Wooldridge, J. M. (2015) 'What Are We Weighting For?', *Journal of Human Resources*. University of Wisconsin Press, 50(2), pp. 301–316. Available at: https://ideas.repec.org/a/uwp/jhriss/v50y2015i2p301-316.html (Accessed: 19 October 2020).

Stokes, J. *et al.* (2015) 'Effectiveness of case management for "at risk" patients in primary care: A systematic review and meta-analysis', *PLoS ONE*. Public Library of Science, 10(7). doi: 10.1371/journal.pone.0132340.

Stokes, J. *et al.* (2017) 'The foundations framework for developing and reporting new models of care for multimorbidity', *Annals of Family Medicine*. Annals of Family Medicine, Inc, 15(6), pp. 570–577. doi: 10.1370/afm.2150.

Stokes, J. *et al.* (2018) 'Towards incentivising integration: A typology of payments for integrated care', *Health Policy*. Elsevier Ireland Ltd, pp. 963–969. doi: 10.1016/j.healthpol.2018.07.003.

Struckmann, V. *et al.* (2017) *How to strengthen financing mechanisms to promote care for people with multimorbidity in Europe? On behalf of the ICARE4EU consortium*, *depositonce.tu-berlin.de*. Available at: https://depositonce.tu-berlin.de/handle/11303/9791 (Accessed: 5 May 2020).

Struijs, J. N. *et al.* (2012) 'Effects of bundled payment on curative health care costs in the Netherlands : An analysis for diabetes care and vascular risk management based on nationwide claim data, 2007-2010'. Rijksinstituut voor Volksgezondheid en Milieu RIVM.

The Academy of Medical Sciences (2018) *Multimorbidity: a priority for global health research*.

Tsiachristas, A. *et al.* (2016) 'Impact of financial agreements in European chronic care on health care expenditure growth', *Health Policy*. Elsevier Ireland Ltd, 120(4), pp. 420–430. doi: 10.1016/j.healthpol.2016.02.012.

Violan, C. *et al.* (2014) 'Prevalence, determinants and patterns of multimorbidity in primary care: A systematic review of observational studies', *PLoS ONE*. Public Library of Science, 9(7). doi: 10.1371/journal.pone.0102149.

Weiss, C. O. *et al.* (2014) 'Multimorbidity and evidence generation', *Journal of General Internal Medicine*. Springer New York LLC, 29(4), pp. 653–660. doi: 10.1007/s11606-013-2660-5.

World Health Organisation (2015) *WHO global strategy on people-centred and integrated health services Interim Report*. Available at: www.who.int (Accessed: 5 May 2020).

World Health Organization (2016) *Multimorbidity*. World Health Organization. Available at: https://apps.who.int/iris/handle/10665/252275 (Accessed: 5 May 2020).

World Health Organization (2020) *Constitution*. Available at: https://www.who.int/about/who-we-are/constitution (Accessed: 28 September 2020).

Wubulihasimu, P., Brouwer, W. and van Baal, P. (2016) 'The Impact of Hospital Payment Schemes on Healthcare and Mortality: Evidence from Hospital Payment Reforms in OECD Countries', *Health Economics (United Kingdom)*. John Wiley and Sons Ltd, 25(8), pp. 1005–1019. doi: 10.1002/hec.3205.

Acknowledgements

This project (SELFIE) has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 634288. The content of this presentation reflects only the SELFIE groups' views and the European Commission is not liable for any use that may be made of the information contained herein.
Jonathan Stokes was additionally supported by an MRC Fellowship.
The SHARE data collection has been funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N°211909, SHARE-LEAP: GA N°227822, SHARE M4: GA N°261982) and Horizon 2020 (SHARE-DEV3: GA N°676536, SERISS: GA N°654221) and by DG Employment, Social Affairs & Inclusion. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C) and from various national funding sources is gratefully acknowledged*(see www.share-project.org).*
ELSA was developed by a team of researchers based at the University College London, NatCen Social Research, and the Institute for Fiscal Studies. The data were collected by NatCen Social Research. The funding is currently provided by the National Institute of Aging (R01AG017644), and a consortium of UK government departments coordinated by the National Institute for Health Research.

Appendices

Appendix I: Multimorbidity dummy variable

We highlighted the 10 chronic conditions from SHARE which overlapped with ELSA. These included; a heart attack including myocardial infarction or coronary thrombosis or any other heart problem including congestive heart failure; high blood pressure or hypertension; a stroke or cerebral vascular disease; diabetes or high blood sugar; chronic lung disease; asthma; arthritis; cancer or malignant tumour, including leukaemia or lymphoma but excluding minor skin cancers; and Alzheimer's disease, dementia, organic brain syndrome, senility or any other serious memory impairment. The other condition was depression and was obtained from the EURODCAT score in SHARE and the CES-D score (4 or more) from ELSA.

Appendix II: Ordered logistic random effects regressions showing the effects of multimorbidity and differential multimorbidity effect under the three payment methods, on self-assessed health

	1	2	3	4	5
MM	0.136***	0.132***	-0.036***	-0.123***	-0.109***
	(0.005)	(0.006)	(0.003)	(0.004)	(0.004)
MM*PFP	0.005	0.005	-0.001	-0.005	-0.004
	(0.006)	(0.006)	(0.002)	(0.006)	(0.005)
MM*PFC	0.005	0.005	-0.001	-0.005	-0.004
	(0.007)	(0.007)	(0.002)	(0.006)	(0.005)
MM*ALL	-0.001	-0.001	0.000	0.001	0.001
	(0.004)	(0.004)	(0.001)	(0.004)	(0.004)

Notes: N=282902. Robust, country-clustered standard errors in parentheses. Results are presented as marginal effects.
* p<0.1; ** p<0.05; *** p<0.01

All models control for a pay-for-performance, pay-for-coordination, and all-inclusive payment main effect dummy, as well as education, gender, age categories, gender interacted with age categories, and country and year fixed effects.
MM=multimorbidity, PFP=pay-for-performance, PFC=pay-for-coordination, ALL=all-inclusive payment.

Appendix III: Individual random effects regressions results showing the effects of multimorbidity and multimorbid groups under the three payment methods, on health and utilisation, excluding Hungary

	Health outcomes			Utilisation outcomes		
	Self-assessed health	Quality of life (CASP)	Limitations with activities of daily living	How often seen or talked to medical doctor last 12 months	How many of these contacts with a GP	Proportion of doctor visits with a GP in the last 12 months
MM	-0.754***	-3.166***	-0.239***	4.390***	2.603***	-0.038***
	(0.011)	(0.176)	(0.018)	(0.340)	(0.307)	(0.008)
MM*PFP	0.001	-0.515	-0.076***	-0.908**	-0.820**	0.005
	(0.046)	(0.321)	(0.024)	(0.438)	(0.403)	(0.012)
MM*PFC	-0.046	0.196	0.014	0.183	0.209	0.019
	(0.035)	(0.308)	(0.036)	(0.468)	(0.378)	(0.012)
MM*ALL	0.009	0.232	0.004	0.872**	-0.100	-0.022
	(0.028)	(0.292)	(0.029)	(0.409)	(0.597)	(0.025)
Observations	279838	217228	273146	176393	121089	99748

Notes: Robust, country-clustered standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01

All models control for a pay-for-performance, pay-for-coordination, and all-inclusive payment main effect dummy, as well as education, gender, age categories, gender interacted with age categories, and country and year fixed effects. All health outcomes are coded such that higher is "better" health. MM=multimorbidity, PFP=pay-for-performance, PFC=pay-for-coordination, ALL=all-inclusive payment.

Appendix IV: Individual random effects regressions results showing the effects of multimorbidity and multimorbid groups under the three payment methods, on health and utilisation, excluding years where payment information is assumed

	Health outcomes			Utilisation outcomes		
	Self-assessed health	Quality of life (CASP)	Limitations with activities of daily living	How often seen or talked to medical doctor last 12 months	How many of these contacts with a GP	Proportion of doctor visits with a GP in the last 12 months
MM	-0.713***	-2.748***	-0.182***	3.873***	2.520***	-0.030***
	(0.018)	(0.201)	(0.021)	(0.373)	(0.303)	(0.009)
MM*PFP	-0.075**	-1.325***	-0.098***	0.370	0.157	-0.008
	(0.035)	(0.468)	(0.036)	(1.462)	(1.179)	(0.014)
MM*PFC	-0.042	0.056	-0.002	-0.016	0.123	0.016
	(0.058)	(0.343)	(0.025)	(0.571)	(0.439)	(0.020)
MM*ALL	-0.025	0.282	-0.013	1.188**	0.155	-0.027
	(0.031)	(0.334)	(0.018)	(0.560)	(0.553)	(0.027)
Observations	110249	63252	104007	76183	76124	65601

Notes: Robust, country-clustered standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01

All models control for a pay-for-performance, pay-for-coordination, and all-inclusive payment main effect dummy, as well as education, gender, age categories, gender interacted with age categories, and country and year fixed effects. All health outcomes are coded such that higher is "better" health. MM=multimorbidity, PFP=pay-for-performance, PFC=pay-for-coordination, ALL=all-inclusive payment.