Won't you stay just a little bit longer? A Discrete Choice Experiment of UK Doctors' Preferences for Delaying Retirement

Abstract

Introduction and aims

Health systems around the world face difficulties retaining their workforce, which is exacerbated by the early retirement of experienced clinicians. This study aims to determine how to incentivise doctors to delay their retirement.

Methods

We used a discrete choice experiment to estimate the relative importance of job characteristics in doctors' willingness to delay retirement, and the number of extra years they were willing to delay retirement when job characteristics improved. 2885 British Medical Association members aged between 50 and 70 years, registered with the General Medical Council, practising in Scotland (in December 2019), and who had not started to draw a pension were invited. We compared the preferences of hospital doctors (HDs) and general practitioners (GPs).

Results

The response rate was 27.4% (n=788). The number of extra years expected to work was the most important job characteristic for both respondents, followed by work intensity for GPs, whereas working hours and on-call were more important for HDs. Personalised working conditions and pension taxation were the least important characteristics for both groups. Setting all characteristics to their BEST levels, GPs would be willing to delay retirement by 4 years and HDs by 7 years.

Conclusions

Characteristics related to the job rather than pension could have the greatest impact on delaying retirement among clinicians.

Introduction

Health systems around the world face difficulties creating and maintaining an effective, efficient and motivated workforce capable of meeting the health needs of growing and ageing populations [1-4]. Workforce issues can be addressed in various ways – increasing training places, improving retention of qualified staff, improving productivity of the existing workforce, and managing demand. In medical education, notable effort is focused on recruiting and training the next generation of healthcare professionals [4-6]. This focus has resulted in the relative neglect of examining how best to retain experienced staff; those nearing the end of their careers who have much tacit knowledge about patient care, education, leadership, research, and innovation [7-9].

Yet the risks to healthcare delivery and medical education posed by the loss of experienced clinicians through retirement are immediate in many countries. For example, in the United Kingdom (UK), two-thirds of doctors aged 55-64 are considering retiring within three years [10] and a large proportion of these doctors plan to take early retirement [11-17]. In the United States (US) the average retirement age of physicians has decreased notably over the last 20-30 years [18-19]. The loss of such experience-based knowledge endangers healthcare delivery and healthcare education in the near future [11-14, 20].

Identifying how best to incentivise doctors to stay in the workforce is context dependent. What is valued in the workplace and associated decisions about working life, such as whether to retire or not, are made against a background of societal norms about working life and retirement age, norms which are typically related to pension eligibility and provision [21-22]. Thereafter, influences on retirement decisions are multi-faceted - individual (e.g., poor health, retirement of a spouse, age-related deterioration) [23-27] and/or job-related (e.g., work autonomy, work flexibility, provision of rewards/recognition, relationships with work colleagues) [28-32]. The complexity of retirement decision-making is compounded by the fact that retirement is associated with both negative – e.g.,

loss of professional identity [e.g., 33] – and also positive factors - related to better health and higher life satisfaction [e.g., 34]. Moreover, this is a dynamic space where national (e.g., pension reforms) and organisational policies and practices change constantly [e.g., 21, 22, 35, 36], and these macro-level changes impact individuals' behaviour [21]. However, to the best of our knowledge, there are no studies directly examining the relative importance of these different factors, and the trade-offs that doctors are willing to make when considering retiring or staying in the workforce.

To address this gap in the literature, the aim of the current study is to determine the relative importance of, and trade-offs between, job-related factors that influence UK doctors' decisions to retire from clinical work. We also estimate whether, by setting these factors at different levels, job conditions could be created that would incentivise doctors to delay their planned retirement and remain in the workforce.

Methods

Design

The study employed the discrete choice experiment (DCE) methodology. This approach has been used to elicit health preferences in a range of areas, including doctors' career preferences [37-43]. DCEs ask individuals to state preferences for hypothetical alternatives, each described by several attributes (job characteristics in this study) and associated levels. By asking individuals to make such choices, we can determine the relative importance of the attributes as well as the trade-offs between them, e.g., how much longer would respondents be willing to work if work intensity reduced.

Recruitment of Participants

Invitations to participate in the study were sent by the British Medical Association (BMA) Scotland, via email, on behalf of the research team. All BMA Scotland members aged between 50 and 70 years who were members of the General Medical Council (GMC) and were practising wholly or mainly in Scotland at the time of the study (December 2019), were invited to take part (n=2885). Age of invitees was restricted to include those most likely to be approaching retirement. Respondents eligible for inclusion were general practitioners (GPs) and hospital doctors (HDs) who had a current clinical role in the National Health Service (NHS) and had not yet started to draw a pension. One reminder email was sent to encourage those who had not participated to do so.

Discrete choice experiment

The first stage of a DCE is to define the attributes and levels. Following best practice, we used published literature and expert opinion to inform the design of a semi-structured interview schedule to explore factors influencing doctors' retirement decisions [37, 38, 44-47]. We then interviewed 40 doctors, representing GPs, consultants and SAS (associate specialist and specialty) doctors aged between 50 and 68 years; for details see Cleland et al [48]. Thematic analysis resulted in 12 main themes with associated sub-themes: Autonomy (perceived degree of clinical autonomy); Facilities (buildings/accommodation, reorganisation, IT); Finance & Employment (pension, taxation, personal finances, 24hr retirement); Management (perceptions of managers, communication, trust/respect); Patients (patient expectations/attitudes); Personal (family circumstances, retirement intentions, plans for later life); Personal development (continuing education, career opportunities); Quality of care (safety, confidence, competence, deskilling); Regulation (attitudes to/experience of appraisal/revalidation); Team/Colleagues (team functioning, communication, trust/respect, staff shortages); Well-being (work-life balance, work-related stress, job satisfaction); and Workload (work intensity, volume of work, working hours, on-call/out-of-hours).

Sub-themes that could not be influenced by policy (e.g. family circumstances) or were a necessary part of the job (e.g. revalidation) were excluded [38, 44]. We then assessed sub-themes most influential among interview participants and compared these with previously published research. Seven attributes, with associated levels, were initially identified from this analysis: *Working hours (described as changes to the current working hours of respondents); On-call/Out-of-hours (i.e. whether the respondents work outside "normal office hours"); Personalised working (i.e. opportunities to personalise respondents' working arrangements that offer more flexibility than standard working arrangements); Work environment* (concerning satisfaction with the workplace)*; Extrinsic job demands (*concerning degree of pressure associated with the job)*; Managerial climate; and Taxation of pension (i.e. changes to tax relief)*. Feedback suggested that *Work environment* and *Extrinsic job demands* were correlated; we combined these into *Work intensity (i.e. degree of pressure respondents experience in their workplace).* We dropped *Managerial climate* as this was likely to mean different things to primary and secondary care doctors. We included an *'Extra years willing to work' (i.e. number of years respondents would be expected to stay in their job beyond their intended retirement age)* attribute, allowing estimation of doctors' willingness to delay (WTD) retirement when attributes improved. Levels for all attributes were determined through consultation with local doctors. Table 1 shows the final set of attributes and levels. The *On-call/Out-of-hours* attribute was excluded from the GPs DCE as most Scottish GPs practices now use centralised out-of-hours services.

Table 1: Attributes and levels for the DCE

Ngene experimental design software was used to identify the choice tasks (49). Using an orthogonal main effects design, and assuming null interaction effects between attributes (i.e. preferences for one attribute did not depend on the level of another), 36 choice tasks were generated. To reduce

cognitive burden, each respondent faced 12 choice tasks; successive respondents received one of three 'blocked' versions, each comprising one-third of the choice tasks.

Respondents were first asked at what age they planned to stop working for the NHS. They were then asked to imagine they were approaching that age; each choice task asked if they would consider continuing to work beyond their intended retirement age. The choice sets were presented as a binary choice: respondents were asked if they would accept the job ('yes' or 'no'). Supplementary Figure S1 provides an example of the DCE choice task and Supplementary Information 1 provides more information on the design and piloting of the DCE. The survey also collected data about respondents' age, gender and health board (which indicated geographical location of work).

During piloting, four additional choice tasks were included; two 'warm-up' questions to familiarise respondents with the task in hand; and one repeated scenario (i.e. two identical choice sets) which served as a consistency check. The vast majority of pilot respondents (n=37) responded in a rational way to the consistency checks, suggesting that they had understood and engaged with the tasks. Given feedback from the pilot work that the questionnaire was long, and to further reduce the cognitive burden, the two consistency tasks and one of the warm-up tasks were omitted from the final survey, making 13 tasks faced in total.

Data analysis

The data was analysed with the random utility maximisation (RUM) framework (50). This assumes that in each choice task ($t = 1, \ldots \ldots, T$), doctors ($n = 1, \ldots . , N$) derive utility (benefit) (V_{ntj}) for a job ($j = 1, \ldots . J$), and choose to take up the job and delay their retirement if the benefit from doing so is greater than not taking the job (and retiring at the planned date). Here we show the deterministic (observed) part of the Utility function, V. The random (unobserved) part is captured by

the error term where we assume errors are independently and identically distributed as type 1 extreme values (IID EV1). We specified a random effects logit regression model. However, models that allow for preference and/or scale heterogeneity were also used to check for robustness. All attributes, other than $EXTRAYEAR$, were modelled as dummy variables, with the BEST level for each attribute as the reference.

The regression equation for HDs is:

$$\begin{aligned} V_{HDs,ntj} = {} & \beta_0 BEST_{ntj} + \beta_1 HOURS2_{ntj} + \beta_2 HOURS3_{ntj} + \beta_3 HOURS4_{ntj} \\ & + \beta_4 ONCALL2_{ntj} + \beta_5 ONCALL3_{ntj} + \beta_6 PERSONAL2_{ntj} \\ & + \beta_7 INTENSITY2_{ntj} + \beta_8 INTENSITY3_{ntj} \\ & + \beta_9 TAXATION2_{ntj} + \beta_{10} TAXATION3_{ntj} \\ & + \beta_{11} TAXATION4_{ntj} + \beta_{12} EXTRAYEAR_{ntj} \end{aligned}$$

And the regression equation for GPs is as follows:

$$\begin{aligned} V_{GPs,ntj} = {} & \beta_0 BEST_{ntj} + \beta_1 HOURS2_{ntj} + \beta_2 HOURS3_{ntj} + \beta_3 HOURS4_{ntj} \\ & + \beta_6 PERSONAL2_{ntj} + \beta_7 INTENSITY2_{ntj} \\ & + \beta_8 INTENSITY3_{ntj} + \beta_9 TAXATION2_{ntj} \\ & + \beta_{10} TAXATION3_{ntj} + \beta_{11} TAXATION4_{ntj} \\ & + \beta_{12} EXTRAYEAR_{ntj} \end{aligned}$$

All regression labels are defined in Table 1.

Results from the regression models were used to calculate the relative importance (RI) of each attribute (Supplementary Information 2). We then estimated respondents' Willingness to Delay (WTD) retirement, calculated as the ratio of the coefficient of interest to the negative of the coefficient on $EXTRAYEAR$. Given the dummy variable coding, with the BEST levels of all attributes as the reference level, the negative of the ratio of β_0 to β_{12} shows the number of years doctors are willing to delay retirement when all attributes are set at their best levels. The ratio β_4 to β_{12} shows how the retirement decision would change if the job offer changed from no on-call/out-of-hours working (BEST) to weekends, daytime working only. Similarly, the ratio β_5 to β_{12} shows how retirement decisions would change if there was a normal share of nights and weekends on-call/out-of-hours.

Results

Respondents' characteristics

Of the 2885 doctors invited to participate, 799 opened the survey. Of these, four did not complete any question and seven were ineligible (not in age range (n=2); not GMC registered (n=5)), representing a 27.4% response rate (788/2878). Of these, 538 respondents had not yet drawn a pension and were eligible for inclusion. GPs mean age was 54.99 (SD 2.88) while HDs mean age was 55.6 (SD 3.36) (p=0.031). HDs, on average, worked more sessions than GPs (9.9 vs 6.6 mean sessions per week).

Using Scottish medical workforce data provided by the Information Services Division (ISD, 2020) we compared our sample of HDs with our target population. Our distribution of HD respondents was similar across Health Boards (p=0.102) but had a smaller proportion of males (52.4% versus 60.6%; p=0.005). There was no data available to do similar analysis for the GPs.

Relative importance of job factors on doctors' decisions to delay retirement

Table 2 shows the Random Effects logit regression results. The results were robust to models allowing for preference and/or scale heterogeneity (results are available from DS on request). The positive constant terms (BEST) for both GPs (2.092) and HDs (2.803) show a general preference to accept the clinical post on offer when all levels of attributes are at their best levels.

For GPs, all five factors had a significant effect on the decision to delay retirement. The β_i parameters all move in the expected direction, providing evidence of the face validity of the results. For example, as we progressively move from the BEST level of taxation to worse levels, the utility (benefit) falls as indicated by the negative signs. Thus, relative to the BEST level of a 100% increase in annual allowance for tax relief (AA), a 50% increase in AA reduces utility by 0.649, a 25% increase

reduces utility by 0.7776, and no change in AA (the worst level) reduces utility by 1.065 compared to the BEST level. As expected, GPs preferred reduced working hours (although a 40% reduction was not statistically significantly different from an 80% reduction), excellent personalised working conditions, mild job intensity, a 100% increase in pension tax AA and fewer additional working years.

The number of extra years expected to work was the most important factor in the decision to delay retirement (Relative Importance (RI)=37%, Figure 1) followed by job intensity (RI=25%). Both these attributes are above the percentage we would expect if all attributes were equally important. Working hours fall slightly below the equi-weight line (RI=17%) with personalised working conditions the least important (RI=8%) followed by taxation of pension (RI=13%).

Table 2: Results of the Random Effects Logit analysis and Relative Importance (RI) of attributes

For HDs, all six attributes influenced their decisions to accept a clinical job after their current planned date of retirement and therefore delay retirement. The β_i parameters again all move in the expected direction, providing evidence of the face validity of the HDs results. The number of extra years expected to work was again the most important factor in the decision to continue working (RI=23%), followed closely by working hours (RI=22%) and on-call commitment (RI=21%) with all three of these attributes being above the equi-weight line. Personalised working conditions were least important (RI=8%, Figure 1) followed by taxation of pension (RI=11%) and job intensity (RI=15%).

Figure 1: Relative importance of attributes – GPs and HDs

Willingness to delay retirement

Given the contractual and working conditions at the time of the survey, the mean number of years GPs in our sample intended to work before leaving the NHS was 5.5 years (SD 3.6). When all factors are set at their BEST levels, GPs would, on average, be willing to delay retirement by a further 4.0 years (2.092/0.524). This is the maximum retirement delay that could be achieved within the parameters of this DCE, its attributes and levels. We can then consider how a movement away from these BEST levels will impact on GPs willingness to delay their retirement using the estimated coefficients from Table 2 as illustrated in Figure 2. For example, with all other factors staying at their BEST levels, but job intensity moving from the BEST level of mild intensity to one that is described as moderately intense, this would result in GPs willing to delay their intended retirement by 2.9 years ((2.092-0.580)/0.524), a reduction of 1.1 years (0.580/0.524) from the maximum achievable of 4 years delay in retirement. With job intensity at its worst level (extremely/very intense) and all other factors at their BEST level, GPs would only be willing to delay retirement by 0.1 years ((2.092-2.057)/0.524), a reduction of 3.9 years (2.057/0.524) from the maximum 4 years achievable. This represents the largest drop when moving from the best to worst levels of any one attribute and reflects the high relative importance of the job intensity attribute to GPs. Similarly, with all other factors at their BEST levels but with the pension AA moving to a lower 50% increase, this would result in GPs willing to delay retirement by 2.8 years ((2.092-0.649)/0.524). And if there was no change in AA (the worst level), but all other factors at their BEST levels, this would result in GPs willing to delay retirement by 2.0 years ((2.092-1.065)/0.524), compared to the maximum 4.0 years achievable under BEST conditions.

Figure 2: Willingness to delay retirement - GPs and HDs

Given the contractual and working conditions at the time of the survey, the mean number of years HDs intended to work before leaving the NHS was 5.2 years (SD 3.6). When all factors are set at their

BEST levels, HDs would, on average, be willing to delay retirement by a further 7.0 years (2.803/0.403). Figure 2 shows how this would be reduced as a consequence of movements away from the BEST levels for all factors individually. For example, if working hours were reduced from 80% (the BEST level) to 40%, with all other factors at their BEST level, HDs would be willing to delay their intended retirement by 5.3 years ((2.803-0.657)/0.403). A 20% reduction in working hours would achieve a willingness to delay retirement of 3.4 years ((2.803-1.413)/0.403) and no change in working hours would result in a willingness to delay retirement of 1.2 years ((2.803-2.326)/0.403), assuming all other factors remain at their BEST levels. This represents the largest drop from the maximum delay in retirement achievable of 7.0 years from best to worst levels of any one attribute and reflects the high relative importance of the working hours attribute to HDs.

Discussion

This is the first study to explore the relative importance of job-related factors that influence UK doctors' (HDs and GPs) decisions to delay retirement. We found differences between HDs and GPs. Work intensity was the most important factor for GPs, whereas working hours and on-call were the most important factors for HDs. Personalised working conditions and pensions taxation were the least important attributes for both groups.

Within our sample HDs, on average, worked more sessions than GPs (9.9 vs 6.6 mean sessions per week). Our results indicate the high relative importance of working hours to HDs. The fact that GPs nearing retirement are already working fewer sessions than their hospital peers suggests that GPs have more control over their pre-retirement working patterns than their HDs peers. The difference between GPs and HDs in terms of later career working patterns and how these may relate to extending time in the workforce merits further exploration.

Setting all attributes at their best levels, GPs and HDs would be willing to delay their retirement plans by four and seven years, respectively. While it is unrealistic to presume that these "ideal" conditions could be met in today's beleaguered NHS, work intensity and working hours (particularly if on-call hours are included in the latter) seem to be most important in UK doctor retirement decisions [37, 40, 48]. These findings reflect previous studies from other occupational groups [33, 51, 52]. They also cast doubt on the popular and medical media assumptions that pension taxation changes are the major influence on doctors' decisions to retire. While UK doctors will no doubt welcome less stringent pension taxation measures, which have just been introduced, interventions that ensure sufficient staff to manage workload and good working conditions are likely to be more effective in encouraging skilled doctors to remain in the workforce than those which address pensions taxation only.

Comparison with previous literature

In our study, GPs were more likely than HDs to cite 'pressure of work' as a reason for retiring, while HDs were most likely to cite working hours and out-of-hours work. These results are consistent with data from surveys of UK doctors of around retiral age [53-59]. For example, Smith et al.'s [53, 59] respondents stated that workload reduction and shorter hours would encourage them to stay in medicine for longer. Similarly, Marchand and Peckham's 2017 [60] systematic review of the evidence for different approaches to retention and recruitment of GPs found working hours and job intensity influenced job dissatisfaction, which in turn was a significant predictor of GP retention and turnover.

This is a dynamic space where national and organisational policies and practices change constantly and influence individuals' behaviour. We compared our findings with those from a DCE with Scottish hospital consultants 20 years previously, before changes in the consultants' contract [61]. In the earlier study, hours of work was the least important attribute to hospital doctors, compared to being one of the most important factors in this current study. We suggest that the reason for this

difference may be related to increased work intensity over time for HDs but this requires further study. Conversely, earlier studies identified that GPs preferred practices with low out-of-hours work commitments [62]. Since this time, changes in arrangements mean Scottish GPs practices now use centralised out-of-hours services: an On-call/Out-of-hours attribute is no longer relevant to this population.

DCEs must be tailored to the research question and context. For example, a recent DCE from China identified that primary care providers value highly welfare (monetary) benefits, availability of equipment and respect from the community [63]. None of these attributes were identified in our development of the current DCE, emphasizing the need for robust development work. Given this, it is difficult to compare across different DCEs. However, there are similarities with work from the Medicine in Australia: Balancing Employment and Life (MABEL) group in Australia. MABEL studies concluded that increasing job satisfaction and reducing working hours are ways of keeping doctors in the clinical workforce for as long as possible [64, 65].

Strengths and weaknesses of the study

Our use of the DCE methodology has extended our understanding of preferences related to retirement. Unlike the usual survey methodology which dominates this field of research [1, 19, 66], a DCE can identify the relative strength, or value of career-related preferences, as well as the trade-offs respondents are willing to make [38, 40, 67]. The DCE attributes were defined through a best-practice process [46, 47, 68] and we believe we have reported sufficient detail in this paper and the supplementary files to allow quality assessment [69]. This study addresses a gap in the literature. Most DCEs examining doctor and other healthcare professionals' preferences for aspects of work focus on those starting out on their careers [39, 41, 42] often with the aim of identifying how to attract/recruit a healthcare workforce. Our focus was how best to retain expertise – how to keep experienced staff who can continue to contribute to healthcare services and patient care and

educate the next generations of medical students and doctors in training - in the workforce. On the other hand, a DCE does not measure actual behaviour – individuals consider hypothetical scenarios/jobs [e.g., 38, 40] – so we do not know if stated preferences about delaying retirement and actual behaviour would diverge.

The DCE focused solely on job-related attributes since no job-focused intervention can control for the influence of personal factors, such as personal health or a spouse that has retired, on retirement decisions. Talking to those who have retired may provide insight into how job- and personal- factors are weighted in retirement decision-making.

A strength of the current study is the comparison between GPs and HDs: many studies on this topic have only looked at one group of doctors, most frequently general practitioners/primary care doctors [66]. Our online survey enabled us to reach a large number of doctors spread over geographical distances and a variety of different workplaces. The response rate was sufficiently high to allow appropriate statistical analysis. However, response bias is possible. Our comparison of characteristics with the target population suggested that whilst the distribution of HD was similar across Health Boards, the proportion of males was lower than for the target population. We thus conducted subgroup analysis for GPs and HDs according to gender and found that the maximum willingness to delay retirement (years) where all attributes are at BEST levels was similar to the pooled results (GPs: pooled = 4.00. females = 3.98. males = 4.05; HDs: pooled = 6.95. females = 6.44. males = 7.48). A limitation of all surveys is that there may be unobservable differences in respondent characteristics e.g. it may motivate the more upset/angry/strongly opinionated to respond.

Implications for research, policy and practice

Differences in healthcare systems, doctors' contracts and working conditions across contexts limit the generalizability of any research [70]. However, comparing across localities would give insight into whether the value placed on certain factors varies by, for example, region or type of working environment (e.g., a large tertiary hospital versus a district general hospital). It would also be worthwhile to investigate the relationship between understaffing in a hospital or practice and retirement intentions, given working hours, job intensity and on-call are likely to be more onerous in such environments. Additionally, different specialties have different physical and cognitive demands [71-73] and so it would be useful to stratify respondents by specialty if possible, to assess if there is an association between taking early retirement or delaying retirement and field of work. A repeated-measures DCE, administered at regular intervals, may also give an indication of how changes in working conditions (positive and negative) impact doctors' intentions to retire.

There are numerous reasons to keep experienced doctors in the workforce. Our main interest is to ensure sufficient doctors are available to deliver care and train the next generations of doctors. However, there is also an economic argument. The real cost of training a healthcare practitioner is often considered as the cost of training (incurred by the individual, governments and clinical workplaces) averaged over expected years in practice. If doctors retire earlier than the time expected by societal norms, the real cost of training increases. If, on the other hand, doctors delay retirement, then the real cost of training decreases [e.g., 74-76]. Thus, rather than focusing all workforce planning efforts on increasing the number of medical students, we suggest that strategies to retain doctors in the workforce could be achieved in a shorter time frame, at a lower cost and ensure the balance between a junior and more experienced workforce.

Our data was collected in December 2019, before the Covid-19 pandemic. Work intensity has increased following the pandemic for already burdened doctors. Many recently retired doctors re-

entered the workforce to help services that were pushed to capacity [e.g., 77]. Data from the UK's British Medical Association (BMA) about doctors' working patterns and quality of working life during the height of Covid-19 suggests that increased workload and burnout, and worries about the backlog of patient care, may be catalysts for early retirement or leaving the NHS. Whether these intentions will be enacted once "normality" resumes remains to be seen. However, these data emphasise the importance of interventions that encourage willingness to stay in active clinical practice in later career years. We suggest that repeating our DCE survey may provide valuable insight into how the pandemic may, or may not, have impacted on what is valued in the workplace and retirement decision making.

Conclusion

In this study, we have tried to untangle the preference of doctors as they consider and plan their retirement, with the view that keeping experienced doctors in the workforce has a positive impact on workforce capacity and the education of the future workforce. Although retirement decisions are based on multiple considerations, there are some clear patterns in our data that can inform policy and practice approaches to retaining experienced doctors in the workforce, and which open up new avenues for future research.

Acknowledgements: Our thanks to those friends and colleagues whose discussions about retirement, lifetime allowances and pensions ceilings were the motivation for this study. Thanks to all those who participated in qualitative interviews and in developing, piloting and completing the survey. Thanks also go to BMA Scotland for the distribution of invitations.

Contributors: JC had the original idea for this project and obtained funding for researcher salaries. All authors contributed to the development of the survey. OE and DS undertook the data management and analysis. MR, TP and JC contributed throughout the analysis and interpretation of the DCE. JC drafted the introduction and discussion, TP drafted the methods, MR drafted the analysis and results. All authors reviewed each successive draft and contributed to the final paper. JC and DS are responsible for the overall content as guarantors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: This study was funded by a grant from the University of Aberdeen Development Trust (UOA Ref: RG14022), and the qualitative data collection (reported separately) was supported by funding from BMA Scotland (UOA Ref: RG14434).

Competing interests: None declared.

Patient consent: Not applicable.

Ethics approval and sponsorship: Approval for this study was granted by the College Ethics Review Board of the School of Life Sciences and Medicine at the University of Aberdeen (Ref: CERB/2018/12/1684), and by R&D departments of NHS Research Scotland in all Scottish Health Board Areas (IRAS: 238076). The study sponsor is the University of Aberdeen.

Data sharing statement: The survey instrument and all available data can be obtained by contacting the corresponding author (JC).

REFERENCES

1. Gorman D. Matching the production of doctors with national needs. Med Educ. 2018;52(1):103-113.

2. Gorman D. Developing health care workforces for uncertain futures. Acad Med. 2015;90(4):400-403.

3. Christensen K, Doblhammer G, Rau R, Vaupel JW. Ageing populations: The challenges ahead. Lancet. 2009;374(9696):1196-1208.

4. Frenk J, Chen L, Bhutta ZA, et al. Health professionals for a new century: Transforming education to strengthen health systems in an interdependent world. Lancet. 2010;376(9756):1923-1958.

5. Chen FM, Bauchner H, Burstin H. A call for outcomes research in medical education. Acad Med. 2004;79(10):955-960.

6. Dauphinee WD. Educators must consider patient outcomes when assessing the impact of clinical training. Med Educ. 2012;46(1):13-20.

7. Leonard-Barton D, Swap W, Barton G. Critical knowledge transfer. 1st ed. Harvard Business Review Press; 2014.

8. Brooke L, Taylor P. Older workers and employment: Managing age relations. Ageing Soc. 2005;25(3):415-429.

9. Buyens D, Van Dijk H, Dewilde T, Vos A. The aging workforce: Perceptions of career ending. J Manage Psychol. 2009;24(2):102-117.

10. General Medical Council (GMC). The state of medical education and practice in the UK: 2020 report. https://www.gmc-uk.org/about/what-we-do-and-why/data-and-research/the-state-of-medical-education-and-practice-in-the-uk Accessed 30 March 2021.

11. Moberly T. More hospital doctors are opting to retire early. BMJ (Online). 2018;362:k3744.

12. Moberly T. Rise in GPs taking early retirement. BMJ. 2018;360:ki367.

13. Shin J, Kim YJ, Kim JK, Lee DE, Moon S, Choe JY, Lee WK, Lee HM, Cho KH. Probability of early retirement among emergency physicians. J Prev Med Public Health. 2018;51(3):154-162.

14. The Health Foundation, The King's Fund, The Nuffield Trust. The health care workforce in england. make or break? https://www.health.org.uk/sites/default/files/upload/publications/2018/The-health-care-workforce-in-England.pdf. Updated 2018. Accessed 30 March 2021.

15. Fletcher E, Abel GA, Anderson R, Richards SH, Salisbury C, Dean SG, Sansom A, Marren FC, Campbell JL. Quitting patient care and career break intentions among general practitioners in south west England: Findings of a census survey of general practitioners. *BMJ Open*. 2017;7(4).

16. Owen K, Hopkins T, Shortland T, Dale J. GP retention in the UK: A worsening crisis. findings from a cross-sectional survey. *BMJ Open*. 2019;9(2).

17. Napier J, Clinch M. Job strain and retirement decisions in UK general practice. *Occup Med*. 2019;69(5):336-341.

18. Wall Street Physician. When should physicians retire? https://www.kevinmd.com/blog/2018/01/when-should-physicians-retire.html . Updated 2018. Accessed 30 March 2021.

19. Petterson SM, Rayburn WF, Liaw WR. When do primary care physicians retire? implications for workforce projections. Ann Fam Med. 2016;14(4):344-349.

20. Adler G, Hilber D. Will the types of jobs being created enable older workers to keep working? J Workplace Behav Health. 2008;23(1-2):71-87.

21. National Research Council. Work, retirement and pensions. In: The National Academies, ed. Preparing for an aging world. the case for cross-national research. National Academies Press (US); 2001.

22. Finnish Centre for Pensions. Retirement ages in different countries. https://www.etk.fi/en/the-pension-system/international-comparison/retirement-ages/. Updated 2019. Accessed 30 March 2021.

23. Adams GA, Beehr TA. Turnover and retirement: A comparison of their similarities and differences. Pers Psychol. 1998;51(3):643-665.

24. Ruhm CJ. Why older Americans stop working. Gerontologist. 1989;29(3):294-299.

25. Schmidt JA, Lee K. Voluntary retirement and organizational turnover intentions: The differential associations with work and non-work commitment constructs. J Bus Psychol. 2008;22(4):297-309.

26. National Institute on Aging. Growing older in America: The health and retirement study. https://www.nia.nih.gov/sites/default/files/2017-06/health_and_retirement_study_0.pdf . Updated 2007. Accessed 30 March 2021.

27. Herrbach O, Mignonac K, Vandenberghe C, Negrini A. Perceived HRM practices, organizational commitment, and voluntary early retirement among late-career managers. Hum Resour Manage. 2009;48(6):895-915.

28. Herrbach O, Mignonac K, Vandenberghe C, Negrini A. Perceived HRM practices, organizational commitment, and voluntary early retirement among late-career managers. Hum Resour Manage. 2009;48(6):895-915.

29. Adams GA, Prescher J, Beehr TA, Lepisto L. Applying work-role attachment theory to retirement decision-making. Int J Aging Hum Dev. 2002;54(2):125-137.

30. Saba T, Guerin G. Extending employment beyond retirement age: The case of health care managers in Quebec. Public Pers Manage. 2005;34(2):195-214.

31. McEvoy GM, Blahna MJ. Engagement or disengagement? older workers and the looming labor shortage. Bus Horiz. 2001;44(5):46-52.

32. von Bonsdorff ME. Age-related differences in reward preferences. Int J Hum Resour Manage. 2011;22(6):1262-1276.

33. Onyura B, Bohnen J, Wasylenki D, et al. Reimagining the self at late-career transitions: How identity threat influences academic physicians' retirement considerations. Acad Med. 2015;90(6):794-801.

34. Guerriero Austrom M, Perkins AJ, Damush TM, Hendrie HC. Predictors of life satisfaction in retired physicians and spouses. Soc Psychiatry Psychiatr Epidemiol. 2003; 38:134-41.

35. National Research Council. Work, retirement and pensions. In: The National Academies, ed. Preparing for an aging world. the case for cross-national research. National Academies Press (US); 2001.

36. First Actuarial. Research into the impact of pensions tax in the NHS – summary. https://www.nhsemployers.org/-/media/Employers/Publications/Reward/Pensions-tax-in-the-NHS---Summary.pdf . Updated 2019. Accessed 30 March 2021.

37. De Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: A review of the literature. Health Econ. 2012;21(2):145-172.

38. Ryan M, Gerard K, Amaya-Amaya M. Using discrete choice experiments to value health and health care. 1st ed. Dordrecht, Netherlands: Springer; 2008.

39. Mandeville KL, Lagarde M, Hanson K. The use of discrete choice experiments to inform health workforce policy: A systematic review. BMC Health Serv Res. 2014;14(1).

40. Cleland J, Porteous T, Skåtun D. What can discrete choice experiments do for you? Med Educ. 2018;52(11):1113-1124.

41. Cleland JA, Johnston P, Watson V, Krucien N, Skatun D. What do UK doctors-in-training value in a post? A discrete choice experiment. . Medical Education. 2016;50:189-202.

42. Cleland JA, Johnston P, Watson V, Krucien N, Skatun D. What do UK medical students value most in their career? A discrete choice experiment. Medical Education. 2017;51:839-851.

43. Scanlan G, Cleland JA, Johnston PJ, Walker K, Krucien N, Skatun D. Location and Support are Critical to Attracting Junior Doctors: A Discrete Choice Experiment. BMJ Open 2018, 8 (3)

44. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: A user's guide. Pharmacoeconomics. 2008;26(8):661-677.

45. Ryan M, Kolstad J, Rockers P, Dolea C. How to conduct a discrete choice experiment for health workforce recruitment and retention in remote and rural areas: A user guide with case studies. CapacityPlus, World Bank and World Health Organization. 2012.

46. Coast J, Horrocks S. Developing attributes and levels for discrete choice experiments using qualitative methods. J Health Serv Res Policy. 2007;12(1):25-30.

47. Vass C, Rigby D, Payne K. The role of qualitative research methods in discrete choice experiments: A systematic review and survey of authors. Med Decis Mak. 2017;37(3):298-313.

48. Cleland J, Porteous T, Ejebu O, Skåtun D. “Should I stay or should I go now?”: A qualitative study of why UK doctors retire. Med Educ. 2020: 54 (9); 54:821–831.

49. ChoiceMetrics (2012) Ngene 1.1.1 User Manual & Reference Guide, Australia.

50. McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka P, ed. Frontiers in econometrics. New York, NY: Academic Press; 1974:105-142. Accessed 7 March 2018.

51. Ruhm CJ. Why older Americans stop working. Gerontologist. 1989;29(3):294-299.

52. Vough HC, Bataille CD, Noh SC, Lee MD. Going off script: how managers make sense of the ending of their careers. J Manage Stud. 2015;52(3):414-440.

53. Smith F, Lachish S, Goldacre MJ, Lambert TW. Factors influencing the decisions of senior UK doctors to retire or remain in medicine: National surveys of the UK-trained medical graduates of 1974 and 1977. BMJ Open. 2017;7(9).

54. Sansom A, Calitri R, Carter M, Campbell J. Understanding quit decisions in primary care: A qualitative study of older GPs. BMJ Open. 2016;6(2).

55. Newton J, Luce A, van T, Zwanenberg Z, Firth-Cozens J. Job dissatisfaction and early retirement: A qualitative study of general practitioners in the northern deanery. Prim Heath Care Res Dev. 2004;5(1):68-76.

56. BMA Consultants Committee. Consultants pension survey. https://www.bma.org.uk/collective-voice/committees/consultants-committee/priorities/consultants-pension-survey. Updated 2018. Accessed February, 2019.

57. Royal College of Physicians. Focus on physicians: 2017–18 census (UK consultants and higher specialty trainees). https://www.rcplondon.ac.uk/projects/outputs/focus-physicians-2017-18-census-uk-consultants-and-higher-specialty-trainees. Updated 2018. Accessed June, 2019.

58. Taylor K, Lambert T, Goldacre M. Future career plans of a cohort of senior doctors working in the national health service. J R Soc Med. 2008;101(4):182-190.

59. Smith F, Goldacre MJ, Lambert TW. Retirement ages of senior UK doctors: National surveys of the medical graduates of 1974 and 1977. BMJ Open. 2018;8(6).

60. Marchand C, Peckham S. Addressing the crisis of GP recruitment and retention: a systematic review. British Journal of General Practice 2017; 67 (657): e227-e237.

61. Ubach C, Scott A, French F, Awramenko M, Needham G. What Do Hospital Consultants Value About Their Jobs? A Discrete Choice Experiment. BMJ 2003; 326: 1432-1435.

62. Scott A. Eliciting GPs' preferences for pecuniary and non-pecuniary job characteristics. Journal of Health Economics. 2001; (20): 329-347.

63. Song K, Scott A, Sivey P, Meng Q. Improving Chinese primary care providers' recruitment and retention: a discrete choice experiment. Health Policy and Planning. 2015; 30: 68-77.

64. Joyce CM, Wang WC, McDonald HM. Retirement patterns of australian doctors aged 65 years and older. Australian Health Review. 2015;39(5):582-587.

65. Vaswani N, Eyre H, Wang W, Joyce C. Characteristics of doctors transitioning to a non-clinical role in the MABEL study. Asia Pacific Journal of Health Management. 2015;10(2):33-41.

66. Silver MP, Hamilton AD, Biswas A, Warrick NI. A systematic review of physician retirement planning. Hum Resources for Health. 2016;14(1).

67. Lancaster KJ. A new approach to consumer theory. Journal of Political Economy. 1966;74(2):132-157.

68. Kløjgaard ME, Bech M, Søgaard R. Designing a stated choice experiment: The value of a qualitative process. J Choice Model. 2012;5(2):1-18.

69. Seokhai V, de Bekker-Grob EW, Ellis AR, Vass CM. Discrete Choice Experiments in Health Economics: Past, Present and Future. PharmacoEconomics. 2019; 37:201–226.

70. Cleland JA, MacLeod A, Ellaway RH. The curious case of case study research. Med Educ. Early online 27 April 2021 https://doi.org/10.1111/medu.14544

71. Thomas LA, Milligan E, Tibble HM. Health, performance and conduct concerns among older doctors: a retrospective cohort study of notifications received by medical regulators in Australia. J Patient Saf Risk Manage. 2018;23(2):54-62.

72. Peisah C, Wilhelm K. Physician don't heal thyself: a descriptive study of impaired older doctors. Int Psychogeriatr. 2007;19(5): 974-984.

73. Kaups KL. Competence not age determines ability to practice: ethical considerations about sensorimotor agility, dexterity, and cognitive capacity. AMA J Ethics. 2016;18(10):1017-1024.

74. Segal L, Marsh C, Heyes R. The real cost of training health professionals in Australia: it costs as much to build a dietician workforce as a dental workforce. Journal of Health Services Research & Policy. 2017;22(2):91-98.

75. Rosenthal J, McKinley RK, Smyth C, Campbell JL. The real costs of teaching medical students in general practice: a cost-collection survey of teaching practices across England. British Journal of General Practice 2020; 70 (690): e71-e77.

76. Regenstein M, Nocella K, Jewers MM, Mullan F. The Cost of Residency Training in Teaching Health Centers. N Engl J Med 2016; 375:612-614.

77. Peisah C, Hockey P, Benbow S, Williams B. Just when I thought I was out, they pull me back in: The older physician in the COVID-19 pandemic. International Psychogeriatrics 2020; 32(10), 1211-1215.

Table 1: Attributes and levels for the DCE

LABEL	DESCRIPTION	REGRESSION LABELS	LEVELS	REGRESSION PARAMETER ($\boldsymbol{\beta_i}$)
WORKING HOURS	This attribute describes possible changes in your **current** working hours.	HOURS	1. 80% decrease in **current** hours[#] 2. 40% decrease in **current** hours 3. 20% decrease in **current** hours 4. No change	 β_1 β_2 β_3
ON-CALL/OUT-OF-HOURS*	This describes your involvement in working outside of "normal office hours".	ONCALL	1. No on-call/out-of-hours working[#] 2. Weekends, daytime working only 3. Normal share of nights and weekends on-call/out-of-hours	 β_4 β_5
PERSONALISED WORKING	This attribute describes the opportunities that would be available to personalise your working arrangements, e.g. annualised hours, unpaid leave.	PERSONALISED	1. Excellent opportunities available[#] 2. Limited opportunities available	 β_6
WORK INTENSITY	This describes the degree of pressure you would feel in your job which might stem from e.g. volume of work, time constraints, performance targets, staffing levels, IT issues, concerns about patient safety etc.	INTENSITY	1. Mildly/not very intense[#] 2. Moderately intense 3. Extremely/very intense	β_7 β_8
TAXATION OF PENSION	This is about potential changes to the amount that you could contribute to your pension each year while still receiving tax relief ("annual allowance") and assumes a commensurate change in the lifetime allowance.	TAXATION	1. 100% increase in annual allowance[#] 2. 50% increase in annual allowance 3. 25% increase in annual allowance 4. No change in annual allowance	β_9 β_{10} β_{11}
EXTRA YEARS	This attribute indicates the number of years you would be expected to stay in the job, beyond the age you think you will stop working for the NHS.	EXTRAYEARS	1. 1 more year[#] 2. 3 more years 3. 5 more years 4. 7 more years	β_{12}

* Hospital doctors only

[#] BEST level for attribute, reference level for regression analysis, captured in $\beta_{0\ \mathrm{BEST}}$

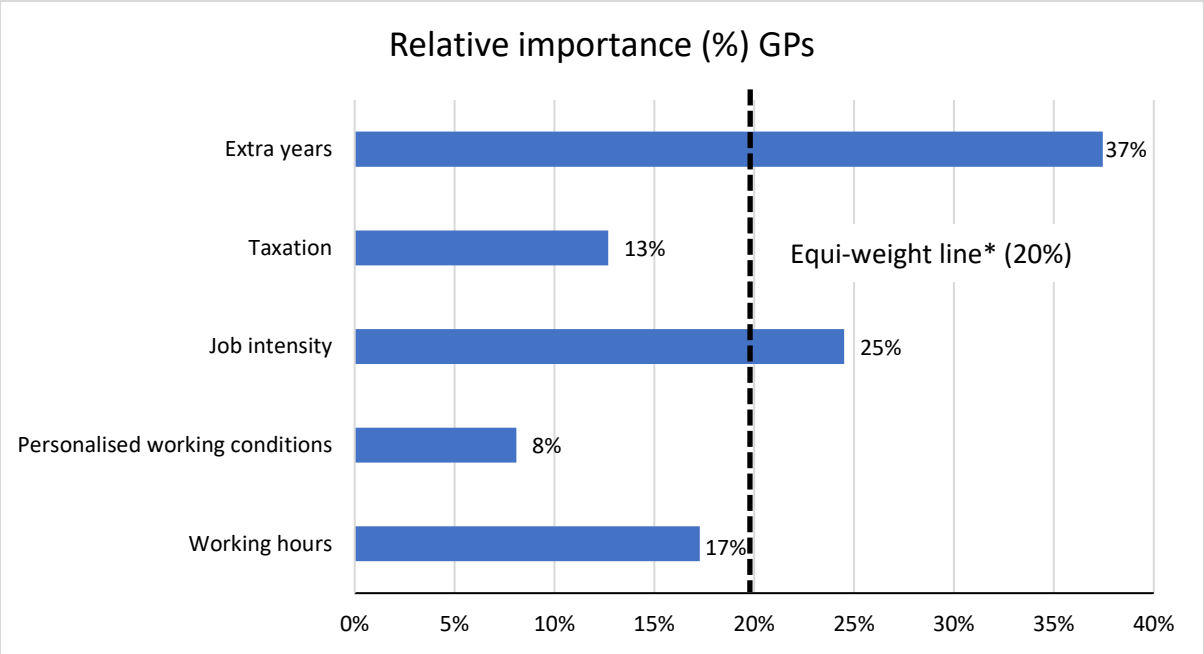
Table 2: Results of the Random Effects Logit analysis and Relative Importance (RI) of attributes

	General Practitioners		Hospital Doctors	
	β_i [Standard Error]	Relative Importance[#] (%)	β_i [Standard Error]	Relative Importance[#] (%)
BEST (β_0)	2.092*** [0.284]		2.803*** [0.274]	
Working hours (Reference: 80% decrease in working hours)[##]				
40% decrease in working hours (β_1)	0.176 [0.185]	17%	-0.657*** [0.179]	22%
20% decrease in working hours (β_2)	-0.628*** [0.202]		-1.413*** [0.278]	
no change in working hours (β_3)	-1.450*** [0.268]		-2.326*** [0.213]	
On-call (Reference: no on-call)[##]				
weekends, daytime working (β_4)	n/a	n/a	-0.847*** [0.128]	21%
normal share of night & weekends (β_5)			-2.162*** [0.174]	
Personalised working conditions (Reference: excellent opportunities)[##]		8%		
limited opportunities (β_6)	-0.679*** [0.174]		-0.830*** [0.169]	8%
Job intensity (Reference: mildly/not very intense)[##]				
moderately intense (β_7)	-0.580*** [0.130]	25%	-0.376*** [0.137]	15%
extremely/very intense (β_8)	-2.057*** [0.265]		-1.559*** [0.158]	
Taxation of pension (Reference: 100% increase in annual allowance)[##]				
50% increase in annual allowance (β_9)	-0.649*** [0.197]	13%	-0.337 [0.228]	11%
25% increase in annual allowance (β_{10})	-0.776*** [0.212]		-0.976*** [0.186]	
no change (β_{11})	-1.065*** [0.222]		-1.207*** [0.154]	
Extra years (β_{12})	-0.524*** [0.052]	37%	-0.403*** [0.035]	23%
Observations	2,609		3,824	
Respondents	218		320	
Block 1	108		113	
Block 2	110		100	
Block3	n/a		107	
Wald chi^2	268.7		408.4	
rho	0.369		0.389	

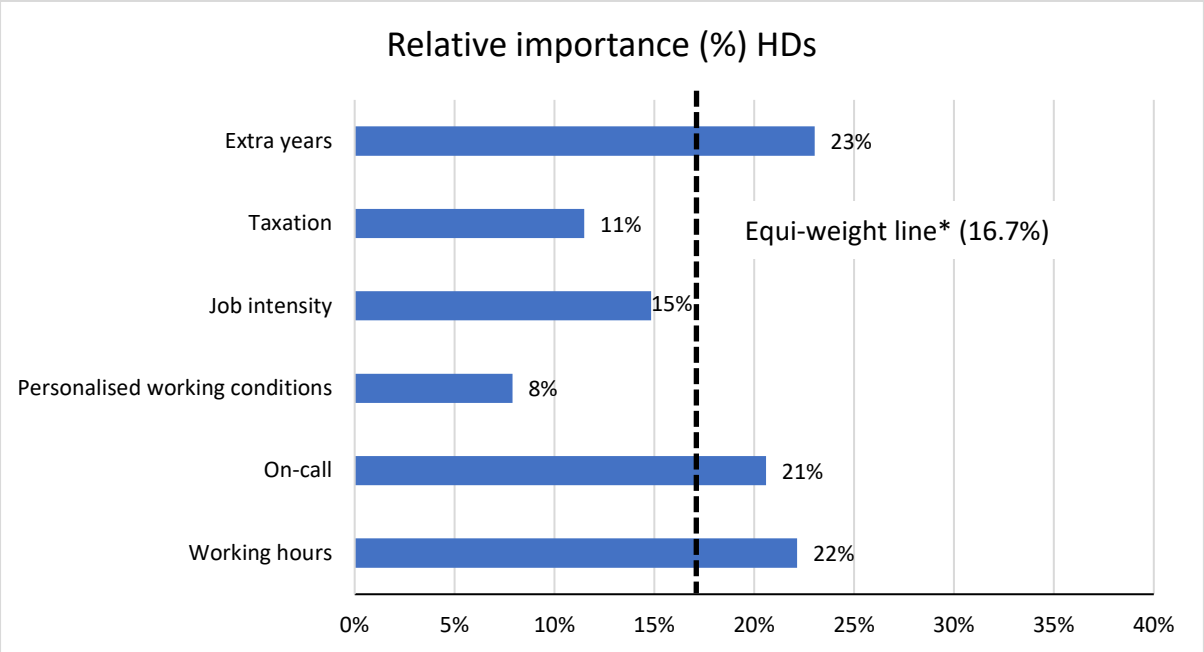
*** p<0.01, ** p<0.05, * p<0.10;

[#] RI of 'working hours' for HDs is calculated as: 2.326/((2.326 +2.162+0.830+1.559+1.207)-(6*-0.403)) = 22.1%;
[##] Variables modelled as dummy variables, with BEST levels as reference levels.

Figure 1: Relative importance of attributes – general practitioners (GPs) and Hospital doctors (HDs)



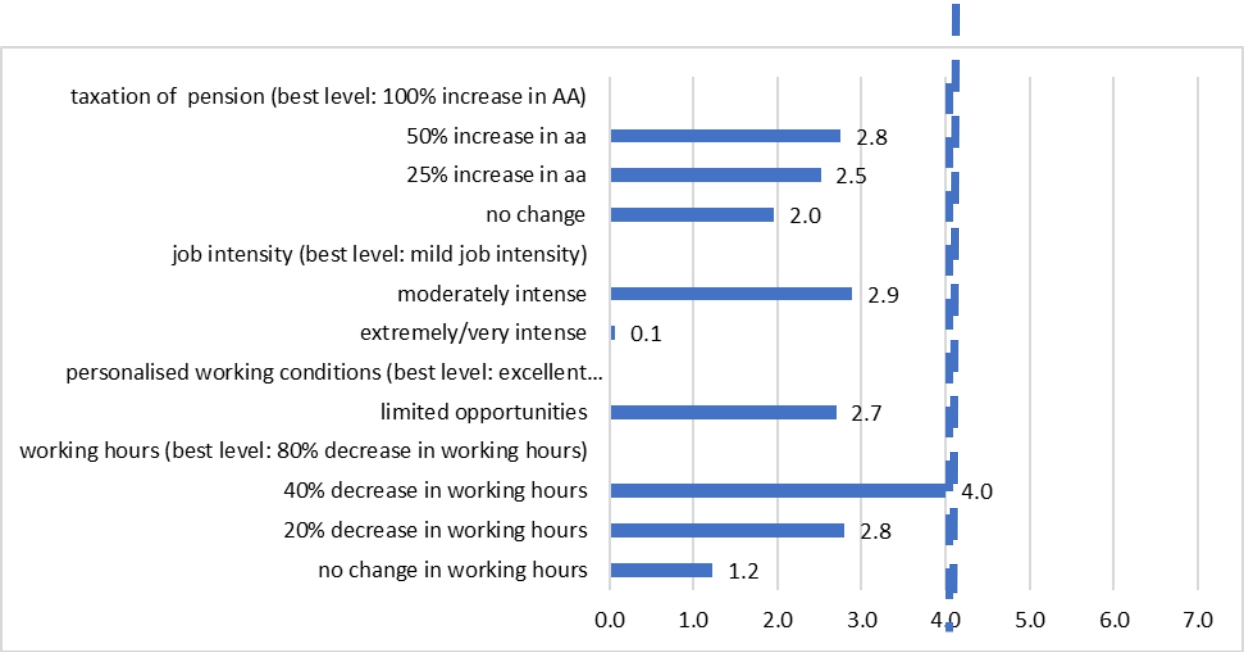
* Equi-weight line indicates that all attributes have equal importance (100/5=20)



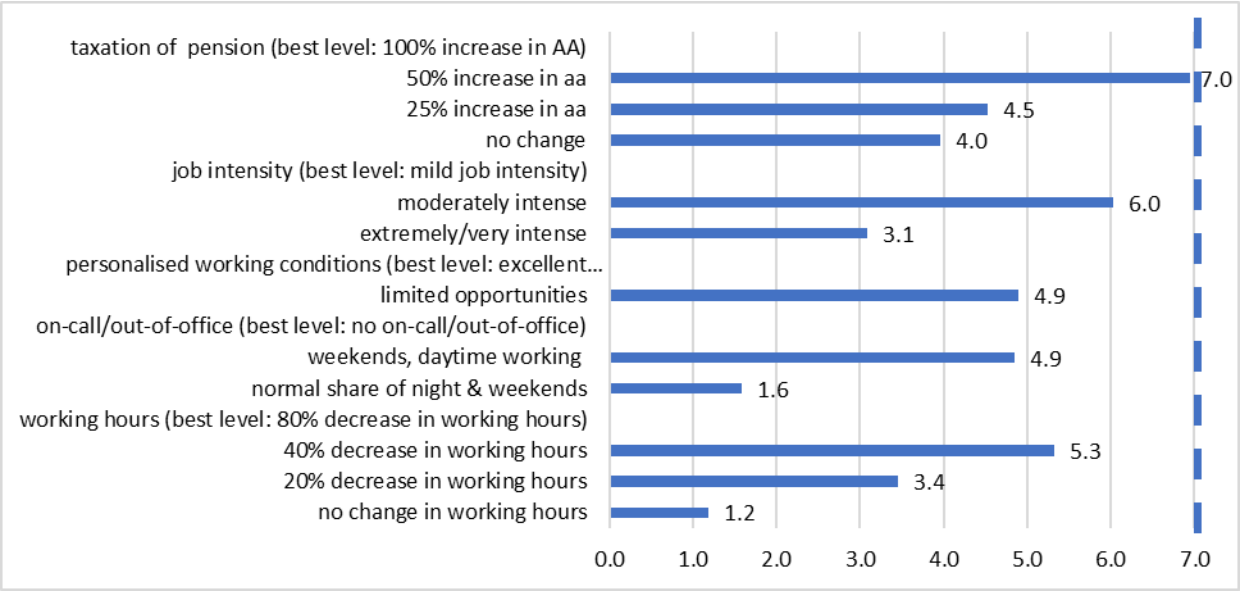
* Equi-weight line indicates that all attributes have equal importance (100/6=16.7)

Figure 2: Willingness to delay retirement (years) as each attribute moves from best to worse levels.

GPs: A maximum of 4 years achievable if all attributes at best level



Hospital Doctors: A maximum of 7 years achievable if all attributes at best level



Supplementary Figure S1: Example of the DCE task

You previously told us you plan to stop working for the NHS when you are X years.

In this section we want to explore the possibility that you might consider continuing to work beyond that age in a clinical role.

- Please imagine you are approaching the age you think you might stop working for the NHS (X years)
- We will present a series of clinical jobs with different characteristics (attributes)
- We will ask you whether you would consider continuing to work beyond X years if you were offered any of these jobs
- In all other respects, the jobs would be the same as your current clinical NHS job (e.g. location, rate of pay, specialty etc)

CHOICE 1:

You are offered a job in your current place of work with the following attributes.

WORKING HOURS	20% decrease in current hours
ON-CALL/OUT-OF-HOURS	Weekends, daytime working only
PERSONALISED WORKING	Limited opportunities available
WORK INTENSITY	Mildly/not very intense
TAXATION OF PENSION	No change in annual allowance
EXTRA YEARS	3 years beyond the age you think you will stop working for the NHS (i.e. X + 3)

Would you accept this job?

Yes ◯ No ◯

Supplementary Information 1: Piloting the DCE survey

An online survey platform (Qualtrics: Provo, UT, USA) was used to script and host the survey, which was piloted in three stages. In *Stage 1*, a convenience sample of local doctors (n=5) was invited to take part in a 'Think aloud' exercise. At a face-to-face interview with OE, participants were asked to complete the online survey and to describe aloud their thinking as they made their choices. Reflecting on the process, they provided feedback, including their overall impressions of the survey, as well as any specific difficulties encountered. This resulted in a few minor refinements to the survey, such as clarifying the language used to describe some terms and questions.

In *Stage 2*, an emailed invitation to complete the revised online survey was sent to participants from the previous qualitative interviews (n=40) (1). Responses from 32 individuals resulted in three main changes. Firstly, some participants commented that the number of choice tasks was excessive. The pilot data indicated that almost all respondents completed the choice tasks consistently. We thus omitted the two consistency choice tasks but kept the warm-up choice task in the final DCE. Secondly, feedback suggested that the *Out-of-Hours* attribute was often irrelevant for Scottish GPs, many of whom have opted out of providing medical services outside normal clinic hours in favour of a centralised out-of-hours service. We dropped this attribute from the GPs survey. In doing so, the experimental design for GPs reduced to 24 choice tasks. We blocked it into two versions, each completed by approximately half of GP respondents. Thirdly, analysis of the pilot data suggested a general preference not to accept the posts offered. We therefore included an 'improved' level for *Working hours* and *Taxation of pension.* The final design for GPs and HDs is shown in Table S1.

Table S1: DCE design

Respondents	**Attributes**	**Levels**	**Full Factorial Design[1]**	**Fractional Factorial Design[2]**	**Blocks**	**Choice questions faced**
Hospital Doctors	6	3 with 4 levels 2 with 3 levels 1 with 2 levels	1,152	36	3	12 + 1 'warm-up'

General Practitioners	5	3 with 4 levels 1 with 3 levels 1 with 2 levels	384	24	2	12 + 1 'warm-up'

1 Full Factorial Design = all possible combinations of attributes and levels
2 Orthogonal Fractional Factorial Design = reduced set of scenarios

In *Stage 3* of piloting, the revised survey was emailed to doctors who had responded positively to our previous invitation to take part in an interview but were not chosen for that stage (n=146) (1). We received 97 full or partial responses and, following data analysis, were content that the DCE and other survey items were working as intended. Only minor modifications to wording were made, meaning that the DCE design did not change and DCE data from this stage could be merged with that from the main Scotland-wide survey.

Supplementary Information 2: Estimating Relative Importance of Attributes

To estimate the relative importance (RI) of a given attribute, the range of that attribute's coefficients (best to worst levels) is divided by the sum of all attributes' ranges and expressed as a percentage. This provides a comparison across attributes through the impact on utility of moving between the extreme levels of each attribute. This can be contrasted with what would be expected if all attributes were valued equally e.g. 100/number of attributes.

For example, assuming $\beta_3, \beta_5, \beta_6, \beta_8$ and β_{11} correspond to the WORST levels for their respective categorical attributes, and β_0 captures the BEST level, and for the continuous attribute, β_{12} corresponds to the range 1 to 7 for the extra years attribute, the RI of $HOURS$ for HDs can be represented by the full formula as:

$$\frac{(\beta_0 - (\beta_0 + \beta_3(1)))}{[(\beta_0 - (\beta_0 + \beta_3(1))) + (\beta_0 - (\beta_0 + \beta_5(1))) + (\beta_0 - (\beta_0 + \beta_6(1))) + (\beta_0 - (\beta_0 + \beta_8(1))) + (\beta_0 - (\beta_0 + \beta_{11}(1))) + (\beta_{12}(1) - \beta_{12}(7)]}$$

This can be simplified and calculated as follows: $\frac{-\beta_3}{(-\beta_3 - \beta_5 - \beta_6 - \beta_8 - \beta_{11} - 6\beta_{12})}$ (2).

References

1. Cleland J, Porteous T, Ejebu O, Skåtun D. "Should I stay or should I go now?": A qualitative study of why UK doctors retire. Med Educ. 2020, 54(9); 821-831
2. Orme B. Getting started with conjoint analysis: Strategies for product design and pricing research. 2nd ed. Madison WI: Research Pub Llc; 2009