This discussion paper has been written by officials of the Welsh Government. Ministers have not had an opportunity to comment on the contents. Exemplifications of changes are provided simply to inform discussion by DSG members. They are not Welsh Government proposals or statements of Government policy for or against changes.

PSS Formulae initial steps - *Older Adults' Residential and Domiciliary care (Example)*

Summary

1. Recap:

- FSG has agreed in principle to the updating of the oldest data in the LG formula whilst maintaining the overall methodology of the formula. This will also involve recalibration of the affected service IBAs.

- DSG has considered the application of this work through the Nursery and Primary Education IBA and concluded that in principle this approach should now be carried forward to other service IBAs.

- DSG also considered that changes, if agreed, should be applied to the settlement as a whole with phasing etc considered with respect to the overall financial impact.

- DSG has recognised that of the other IBAs, the social services IBAs are the most significant in terms of revenue distributed and also the most complex to undertake. We agreed we should therefore turn next to those.

- Additional resource will be needed for both data collection and analysis. Welsh Government has identified additional internal resource and will continue to explore securing further capacity within the constraints on budgets etc. The Local Government Data Unit has identified capacity to undertake the data collection in respect of social services IBAs.

Next Steps

2. This paper looks back at the methodology and associated data collection when the Personal Social Services (PSS) Formulae were updated in 2004 following the Review of Standard Spending Assessments (SSAs) in 2000-01 by Pion Economics/ Swansea University. The formulae were further reviewed in 2016.

3. It then summarises the potential way forward in the data collection and modelling the recalibration of these Formulae, beginning with the older adults' residential and domiciliary care element methodology explained as an example.

4. The '**Older adults' residential and domiciliary care'** formula distributes around **£822 million**. This accounts for 10.7% of the overall SSA of £7,716 million.

Older adults' residential and domiciliary care

0.320 x Pensioners with limiting long-term illness

0.303 x Pensioners living alone in households

0.202 x Population aged 85 and over

0.146 x Pension credit claimants aged 65 and over

0.016 x Dispersion threshold 300 (2001)

0.013 x Dispersion threshold 7,500 (2001)

---

5. The **'Children and Young Persons'** formula distributes around **£755 million**.

Children and Young Persons

0.685 x Dependent children in out of work families

0132 x Dependent children in social rented housing

0.097 x Dependent children in overcrowded housing

0.054 x Population aged under 18 in wards with weighted density greater than the Welsh average

0.032 x Dispersion threshold 300 (2001)

---

6. The **'Younger adults' personal social services'** formula distributes around **£777 million**.

Younger adults' personal social services

0.601 x Population aged 18 to 64

0172 x Severe disablement allowance or disability living allowance claimants or personal independence payment, aged 18 to 64

0.157 x Households where head is aged 18 to 64 with no carer

0.040 x Income support, job seekers allowance, pension credit or universal credit (not in employment) claimants, aged 18 to 64

0.013 x Adults age 18 to 64 in non-white ethnic groups

0.014 x Dispersion threshold 7,500 (2001)

0.003 x Dispersion threshold 300 (2001)

Section 1: Background and Current Situation

7. The Review of Standard Spending Assessments (SSAs) in 2000-01 by Pion Economics/ Swansea University recommended that hierarchical modelling

should be used in the area of personal social services to better identify the factors driving the need to spend. In order to undertake this work, information would be required at a lower geographical level than the authority as a whole. A further disaggregation of clients into various categories would also be required to more accurately identify spending patterns and needs.

8. The review therefore recommended that small area data be investigated as a possible data source in the development of the SSA formula for Older Adults (OA PSS). Consultants from York University were asked to scope the possibility of collecting the data and to advise on the appropriate methodology. Local authorities, the Local Government and Housing Committee and the Distribution Sub-Group of the Consultative Forum on Finance all considered the consultation document.

9. As part of the review the WG Statistical Directorate undertook a brief survey of Welsh authorities in to ascertain what data was currently available in the early 2000s. Only a handful of authorities responded that they had low-level data available for each of the 3 main categories of older adult care, children's and adult social services. In collecting the data from the authorities it was discovered that a large amount of work was needed on the available data for it to be of use. For example the level of postcoding was often quite low. Information was not generally available on the category of clients.

10. DSG and FSG recommended that work should be undertaken to collect the data from authorities in order to develop the PSS formulae. **Annex A** shows the broad details of the project.

11. The data collection at that time showed there was a large variation in the cost of the provision of services to older adults across local authorities in Wales. Such variations arise for a number of reasons including variations in the type of care, variations in local authority policy, variations in efficiency and variations in accounting practice.

12. The work carried out by York University highlighted a range of factors likely to affect the demand for social services amongst older adults. Factors treated as 'universal' which were identified as important factors in deciding the level of provision of social services are as follows:

- Demography
- Dependency
- Disability

Further 'contingent' factors highlighted as affecting the level of provision as well as the provider are:

- Living arrangements
- Availability of informal carers
- Capacity to pay

13. The current formula is based upon four indicators that carry weights to reflect their importance as proxies for the relative demand for service utilisation and two sparsity indicators.

Section 4: Estimating small area costs (<u>Dependent variable)</u>

Database of older adult clients 2003 Data collection

14. An essential recommendation of the York University team was the compilation of a cost database of older people that could be mapped back to the pre-care electoral division by postcode. In past studies, the difficulty in obtaining data of this type has been well documented. .

15. The experience in 2003 was that while the level of records with postcodes was relatively high in most authorities, this was not the case in 3 authorities. At the time the York University team considered that as long as levels of postcoding were 80% or more this should not introduce any bias in the analysis provided that those clients were evenly distributed across the electoral divisions within authorities. Annex A details the process carried out in 2003.

16. For our current project we hope that the increasing use of electronic records and database systems by local authorities will have made it easier to extract client records. However we will not know the level of ward specific information until we undertake the data collection, The standards for processing personal data have changed significantly since 2003 and we will need to consider the best way to approach this.

Proposed Data Collection Exercise for the PSS formula

17. There are two types of data that has to be collected in the area. The first type of data that we require is the individual client data that is held on local authority systems. In 2003 this data had been collected through a simple upload onto a spreadsheet under columns within the attachment (Annex B). This reflected the number of open cases. Data was supplied via postcodes and handled in accordance with the 1998 Data Protection Act with special attention paid to the technical and organisational measures dealing with security.

18. The second part of the data collection exercise in this area relates to financial data. The aim is to estimate costs that relate to three case categories of older adult persons care:

    - Clients in residential homes

    - Clients in nursing homes

    - Other clients not in care homes

19. The intention is to allocate the costs of older adult persons back to their ward/electoral division of origin and come up with an average cost per older adult in that ward/electoral division. This will act as the dependent variable in the analysis. The diagram below shows this in schematic form.

Schematic approach to the model building exercise

1) Data you supply provides the "needs proxy". Mapping client numbers and unit costs to wards via postcodes creates small area costs.
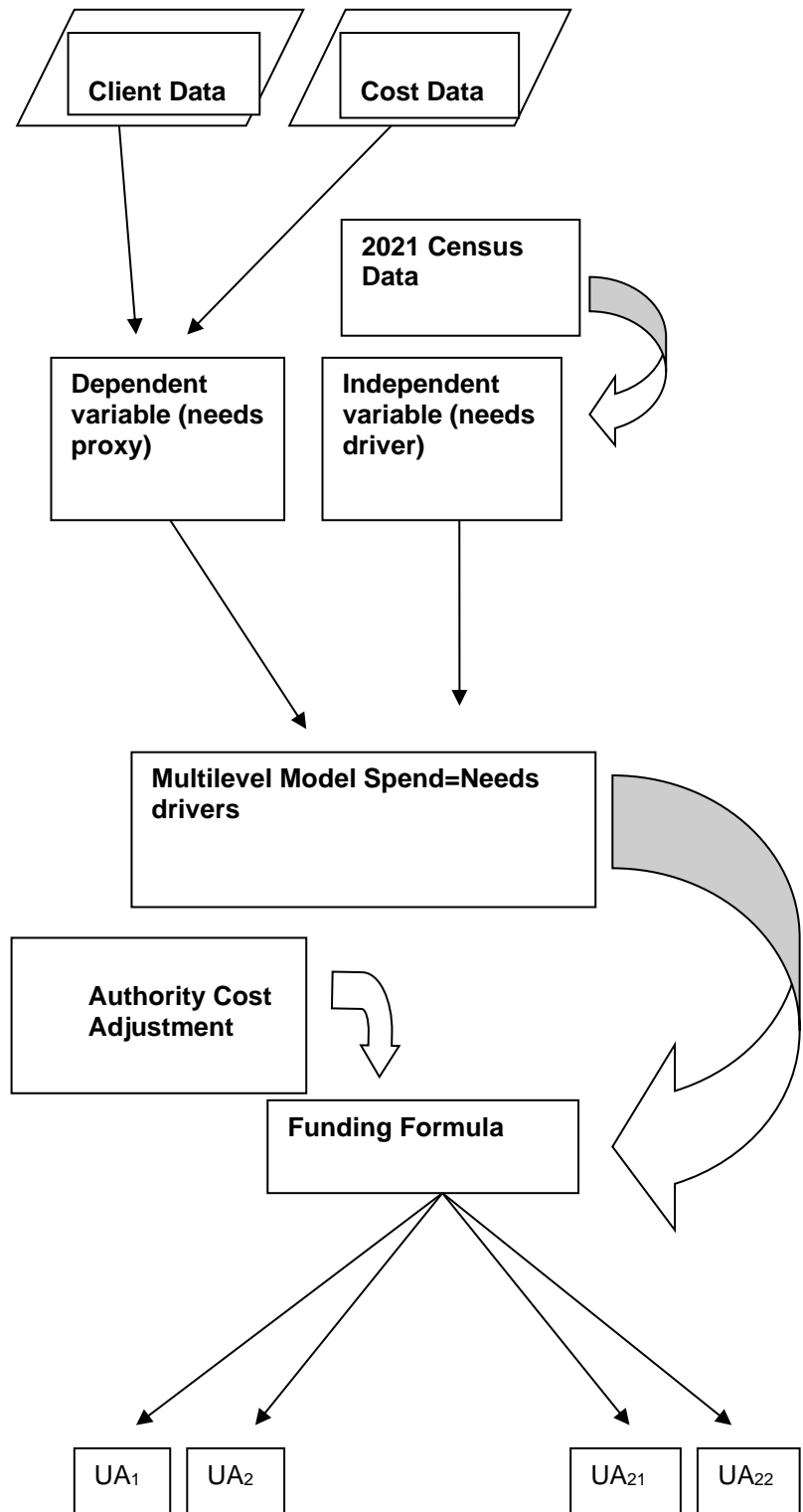
2) Most of the candidate indicators that exhibit strong *theoretical* evidence that they drive need will come from the census.

3) Candidate driver indicators are tested against "need to spend".

4) Candidate drivers most strongly associated with need between wards (1 to i) within authorities (1 to j) are included in the model with suitable weights. The model is concerned with intra-authority variation. Variation from one authority to the next, whether this is due to policy or efficiency, is ignored.

5) If there is strong evidence of authority-level effects on the cost-of-service provision then these may be incorporated into the model.

6) The model should provide a strong evidence-base for updating the funding formula.

| Client Data | Cost Data |

2021 Census Data

| Dependent variable (needs proxy) | Independent variable (needs driver) |

Multilevel Model Spend=Needs drivers

Authority Cost Adjustment

Funding Formula

| UA_1 | UA_2 | .... | UA_{21} | UA_{22} |

Expenditure Need Indicators and Regression Methods

Older adults' residential and domiciliary care	Dependent variable: - Electoral/Ward costs per capita (per 65 year old) Independent variables to be considered (ward level): - Population Aged 65 Plus - Population Aged 75 – 84 - Population Aged 85 Plus - Pensioners with limiting long-term illness - Pensioners living alone in households - Pension credit claimants aged 65 and over - Dispersion threshold 300 (2001) - Dispersion threshold 7,500 (2001) - etc	Unit of Observation Wards Regression Method Hierarchical Modelling

Section 5: Datasets Overview (Independent Variables)

Client Based Data:

- **Population Aged 65 plus**

- **Population Aged 75 – 84**

- **Population Aged 85 Plus**

20. The population data is published by the Office for National Statistics. The latest data uses the latest 2022 mid-year population estimates. Population at Ward level for age groups 65 plus, 75-84 and 85 plus can be built here.

Deprivation Based Data:

- **Pensioners with limiting long-term illness**

- **Pensioners living alone in households**

- **Pension credit claimants aged 65 and over**

- **Etc**

21. The deprivation-based data comprise of mainly Census based data and DWP data. These can be produced using the ONS [custom dataset builder](#) and [NOMIS.](#)

Sparsity Based Data:

- **Dispersion threshold 300 (2001)**

- **Dispersion threshold 7,500 (2001)**

- **Etc**

22. The sparsity data in 2015-16 looked at dispersion data as well as travel speeds and times. **This will be considered in the next stages of the analysis**.

Section 6: Methodology multi-level modelling (MLM)

23. **Multilevel modelling (MLM)** looks at relationships between variables whilst taking into account differences between hierarchical levels in the data. Multi-level small area' modelling employs information on areas *within* each authority and analyses these as well as authority level effects. This adds an extra level to the modelling – hence the term 'multi-level'.

24. Different local authorities will have different levels of service provision. Some of these differences are due to differences in levels of need between the authorities; but some of the differences are due to differences in policies and levels of efficiency between authorities. The resource allocation formula should, as far as is possible, reflect the former rather than the latter.

25. Multi-level modelling is based on the principle that the way that local authorities allocate resources to different areas within their authorities is the best reflection of how spending should respond to need.

26. The application of multi-level modelling can be illustrated with reference to a diagram (Figure 1). In this example there are three local authorities. Each data point relates to a ward within the authority. Need is measured using selected indicators from Census or similar data.

27. In the Personal Social Services IBA, the model analyses the relationship between Older adults' residential and domiciliary electoral ward costs per capita (dependent variable) and population, Pensioners with limiting long-term illness, Pensioners living alone in households and Pension credit claimants aged 65 and over (independent variables) across all authorities, but also takes into account the differences between wards in each authority. The method will test a range of other variables as well as those listed above.

Figure 1
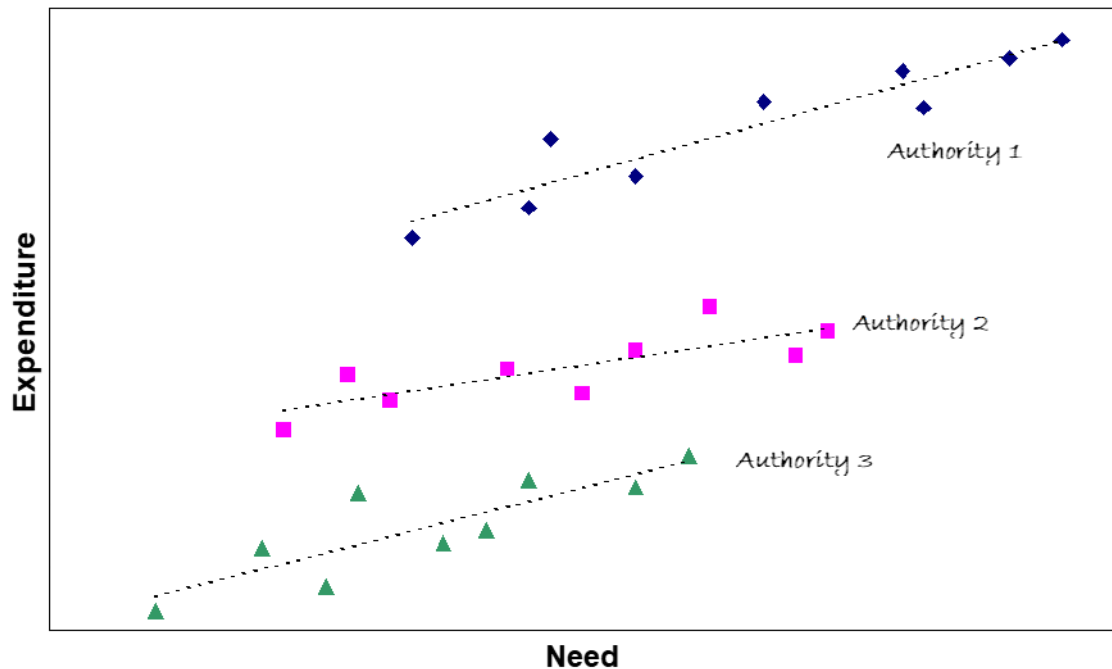
28. The dotted lines show that, as need increases, each authority increases expenditure by an approximately similar amount (shown by the similar slope to the lines). However, authority 1 devotes a higher level of resources overall to the services than authority 2, which in turn devotes more than authority 3.

29. If we are searching for some "standard" response to needs, we should be seeking to identify the average of the individual slopes of the sort given for authority 1, authority 2 and authority 3.

30. The purpose of multi-level modelling is, therefore, to use small area information to arrive at a 'standard' relationship between need for services and expenditure that effectively ignores the effect of different policy decisions between authorities.

31. To arrive at this relationship a modelling strategy will be adopted that combines expert and local judgement with objective statistical criteria. Modelling techniques will be used to select which variables, from a long list of candidates, are included in the formula. The choice will be constrained by a set of prior judgements on the relative importance of a wide range of candidate variables. These judgements will be used to decide which variables are included and excluded from the modelling when the statistical criteria are ambiguous or the results are implausible. They help to ensure that the final formulae are intuitively plausible as well as statistically valid.

Section 7: Discussion and Next Steps

32. **Annex C** provides a rough time timeline based on the investigation of the processes undertaken in the past and the availability of resource. At this stage

we do not know whether it is realistic. New approaches to data recording and extraction may allow for a shorter timetable (or not) but equally capacity in local authorities and in Welsh Government is constrained. DSG is asked to consider this timetable.

33. Alongside the paper DSG have been sent the data collection forms carried out in 2003 for information. In the next few weeks **WG, WLGA and Data Cymru** will meet to discuss the data forms and the variables required that will inform us of the expenditure and ward level for each type of care. **Annex D** provides an example of the final dataset we require to complete the multi-level modelling analysis. **Step 1a** and **Step b** provides the initial steps to the process.

34. DSG are asked to consider and agree the initial steps for the data collection and if possible, advise on **accessibility**, **sensitivity** and **feasibility** of the data collected previously and whether there are more suitable options.

Local Government Finance & Performance
Welsh Government

Project to collect data for 2003 review of PSS formulae

Introduction and summary

The main objectives of the proposed study are to facilitate the National Assembly in developing a new set of formulae and weights for the various Personal Social Services SSAs, making use of small area geographic data wherever possible and appropriate. The aims are: to take forward the recommendations of the independent review; to advise on statistical methodology; to prepare data definitions and make recommendations on the timing and implementation of data collection; to advise on the incorporation of the Learning Disabilities fund into the LGS; and to ensure that the project works in collaboration with the SSIW programme for the development of PMI.

Main activities

More specifically, the project will consist of four main activities.

(1) To recommend modelling methodology for developing allocation formulae for the SSAs for each of the areas of PSS.

These recommendations will be based on a review of methods used for resource allocation of social and health services and the procedures used for resource allocation for services that may substantively or methodologically resemble PSS in the UK and elsewhere. Bearing in mind the discussions in the University of Wales Swansea and Pion Economics Report (hereafter "The Review"), and the subsequent comments of the DSG, we will be concentrating on the feasibility of using multi-level small area modelling to generate allocation formulae.

A particular issue that was raised in The Review was the question of when it is appropriate to use multi-level modelling techniques either because of the structure of the data, or in order to control for variations in local policy and supply. Where data exhibit a very clear hierarchy, such as where small areas are nested within local authorities, then some statistical procedure - such as a fixed effects approach or multi-level modelling - has to be adopted because the assumptions of classical least squares are invalidated.

The project will deal with this question as well as a number of theoretical and practical considerations that are central to the development of evidence based allocation formulae.

These include:

- the feasibility of using small area data;
- the extent of post-coding in client registers and administrative data systems;
- the difficulties of controlling for supply-side factors
- how to deal with multicollinearity amongst likely needs drivers;
- how to deal with determinants of local cost variations in supplying services - notably - how to incorporate weightings or variables to take account of cost differences due to population sparsity;
- how to address major theoretical issues with policy implications - such as the question of whether to base the cost variables in the modelling on gross or net expenditure.

Again, it is worth noting that methodological options cannot be considered in a vacuum, as they critically depend on the availability of suitable data. There are three key practical issues: the nature of the data collection and storage procedures; the associated software and hardware; the data definitions that are in-built in existing and proposed systems; and the local commitment to recording data - which may depend on the perceived relevance of these data to client care or other local uses. Much of the rest of this paper describes how these issues will be tackled.

(2) To recommend methods for computing small area estimates of estimates of service activity and costs - or other types of variables to be used as the dependents in the modelling exercise.

Two main types of data are required to compute such cost estimates: data on client numbers and services received; and data on the costs of services.

The project will assess the availability of these two types of data, both in existing data systems and in the systems that may be installed as a result of proposals from the SSIW and other forthcoming initiatives. It will also need to look at data completeness and quality. After this data scoping exercise, the project will produce recommendations and specification for the data that needs to be collected in order to produce suitable small area cost or activity estimates for the modelling.

These recommendations and specifications need to be sufficiently detailed so that data definitions, collection procedures and quality assurance mechanisms can be built into the current and projected development of local authority computer systems. The project will have a limited role to assist in overseeing and facilitating the early stages of data collection, data entry and quality assurance.

The project will need to ensure that the definitions and any proposals for new systems and procedures are consistent with the programme for the introduction of performance management information currently undertaken by the SSIW.

Depending on the timetable for the introduction of new formulae, and the state of the present data systems, it may be necessary to recommend interim strategies, or less radical options than the creation of entire new data sets. Options to consider include the cost of cleaning and otherwise improving existing data; of conducting small-scale supplementary data collection exercises, including surveys; and the inclusion of comparable data from elsewhere in the UK to bolster the Welsh data in the modelling.

(3) To review existing formula and other sources in order to compile lists of potential needs drivers for each of the PSS sectors.

There are two main possible approaches to identifying an appropriate specification of needs drivers and their weights:

- a quasi-theoretical approach based on a review of literature in order to identify an appropriate candidate set of starting variables, and then searching statistically for a parsimonious sub-set; or
- a normative approach drawing on knowledge and/or experience of providers or eventually clients either only for the selection of variables or both the selection and weighting of variables.

The Review adopted, mainly, the latter approach (see the lists of variables and associated discussion in section 7 of the Review): we would want to extend this by referring to the extant literature in these areas.

12

(4) To check the availability (to small area or other bases) of the variables to be considered as needs drivers.

Small area modelling for resource allocation currently relies heavily on census data for its needs drivers, but this is reducing with the increased availability of small-area administrative data, such as counts of income support and other groups of claimants, and information from police and education authorities.

The project will consider such new data sources as potential needs drivers, and assess their likely availability to small area bases (and their reliability) at the time they will be needed for the modelling.

Nonetheless, the project cannot ignore the potential of census data, especially, if the first modelling does not take place until 2003: by which time small area results should be available for the 2001 census. Moreover, the proposal for more flexible output areas may make the census more attractive for small area work, with: though the precise implications need to be explored.

UNED DDATA LLYWODRAETH LEOL ~ CYMRU
LOCAL GOVERNMENT DATA UNIT ~ WALES

Elderly People Social Services client data (aged 65 and over)

Unitary Authority: _____   Contact name: _____

Tel: _____

Client Data (a)

Date of birth	Gender F/M	Full Postcode (b)	Primary type of care: D - Domiciliary R - Residential N - Nursing O - Other	LA - Local Authority PR - Formerly Preserved Rights RA - Residential Allowance BL - Boyd loophole O - Other	Care home postcode (c)	Client contribuition to care costs per week (£s) (d)	Hours of home help per week (e)
Example…..							
11/12/25	*F*	*PC1 3AZ*	*D*	*LA*	*-*	*-*	*4*
1/21/30	*M*	*PC3 5AJ*	*N*	*PR*	*PC15 1ZE*	*£ 50*	*-*

Proposed timetable for the Personal Social Services (PSS) Formula

Date	Activity
Summer 2024	Agree initial timetable for introducing the formulae.
Summer/ Autumn 2024	Agree expenditure data specification for children and young person's services. (Dependent Variable)
	Agree expenditure data specification for young adults services. (Dependent Variable)
	Agree expenditure data specification for Older adult care services. (Dependent Variable)
Summer/ Autumn 2024	Agree client and deprivation need data specification for children and young person's services. (Independent Variable)
	Agree client and deprivation need data specification for young adults services. (Independent Variable)
	Agree client and deprivation need data specification older adult care services. (Independent Variable)
2024/2025	Collect Independent Variable level data at local authority level and ward level. Client Based Data (Population) and Deprivation Data (Census Indicator Data etc).
Summer/ Autumn 2024	**'Older adults' residential and domiciliary care' formula** Carry out Pearsons' correlation to analyse the best relationship between each of the variables to determine most suitable fit. Carry out multi-level modelling (MLM) on the original 2003 Data using Rstudio to replicate existing analysis.
Spring 2025 Onwards	Start collecting, evaluating and cleaning data for older adult care services. Expenditure data at Ward level.
	Start collecting, evaluating and cleaning data for children and young person's services. Expenditure data at Ward level.
	Start collecting, evaluating and cleaning data for services for young adults services. Expenditure data at Ward level.

Summer/ Autumn 2025	'**Older adults' residential and domiciliary care' formula** Carry out Pearsons' correlation to analyse the best relationship between each of the variables to determine most suitable fit. Carry out multi-level modelling (MLM) using the 2024 Data using the Rstudio set up in the previously.
Summer/ Autumn 2026	'**Older adults' residential and domiciliary care' formula** Determine how to calculate the sparsity element and calculate the weights.
Autumn 2026	DSG Paper exemplifications of options for the 'Older adults' residential and domiciliary care' formula.
Summer/ Autumn 2027	**Children and young person's formula** **young adults formula** Carry out Pearsons' correlation to analyse the best relationship between each of the variables to determine most suitable fit. Carry out multi-level modelling (MLM) using the 2024 Data using the Rstudio set up in the previously.
Summer/ Autumn 2028	**Children and young person's formula** **young adults formula** Determine how to calculate the sparsity element and calculate the weights.
Autumn 2028	DSG Paper exemplifications of options for the **Children and young person's formula**
	DSG Paper exemplifications of options for the **young adults formula**
Summer/ Autumn 2029	Present results of modelling to local authorities and agree models for inclusion in the DSG Report.
Autumn 2029	Incorporate new models in the 2029-30 settlement.

Proposed steps for the Personal Social Services (PSS) Formula – Older adults' residential and domiciliary care' formula

Step 1a: Data Collection (Dependent Variable)

				Residential Costs	Domicilary Costs
			WALES/CYMRU		
00NAMA	NAMA	512Aberffr	512 Aberffraw		
00NAMB	NAMB	512Amlwcl	512 Amlwch Port		
00NAMC	NAMC	512Amlwcl	512 Amlwch Rural		
00NAMD	NAMD	512Beaum	512 Beaumaris		
00NAME	NAME	512Bodffo	512 Bodffordd		
00NAMF	NAMF	512Bodorg	512 Bodorgan		
00NAMG	NAMG	512Braint	512 Braint		
00NAMH	NAMH	512Bryngw	512 Bryngwran		
00NAMJ	NAMJ	512Brynte	512 Brynteg	DATA	
00NAMK	NAMK	512Cadna	512 Cadnant		
00NAML	NAML	512Cefni	512 Cefni		
00NAMM	NAMM	512Cwm C	512 Cwm Cadnant		
00NAMN	NAMN	512Cynga	512 Cyngar		
00NAMP	NAMP	512Gwyng	512 Gwyngyll		
00NAMQ	NAMQ	512Holyhe	512 Holyhead Town		
00NAMR	NAMR	512Kingsla	512 Kingsland		
00NAMS	NAMS	512Llanba	512 Llanbadrig		
00NAMT	NAMT	512Llanbe	512 Llanbedrgoch		
00NAMU	NAMU	512Llanddy	512 Llanddyfnan		
00NAMW	NAMW	512Llaneili	512 Llaneilian		
Etc	Etc	Etc	Etc	Etc	
Etc	Etc	Etc	Etc	Etc	

Step 1b:

				Pop 65+	Ward costs	Pop 74-85	Pop 85+	No car	LLTI	No care	No quals	Living alone	Overcrowding	Social rented	Private rented	No central heating	No basic amenities	Not in self-containe	Inactive	Etc
			WALES/CYMRU																	
NAMA	512Aberffr		512 Aberffraw																	
NAMB	512Amlwcl		512 Amlwch Port																	
NAMC	512Amlwcl		512 Amlwch Rural																	
NAMD	512Beaum		512 Beaumaris																	
NAME	512Bodffo		512 Bodffordd																	
NAMF	512Bodorg		512 Bodorgan																	
NAMG	512Braint		512 Braint						DATA											
NAMH	512Bryngv		512 Bryngwran																	
NAMJ	512Brynte		512 Brynteg																	
NAMK	512Cadna		512 Cadnant																	
NAML	512Cefni		512 Cefni																	
NAMM	512Cwm C		512 Cwm Cadnant																	
NAMN	512Cyngar		512 Cyngar																	
NAMP	512Gwyng		512 Gwyngyll																	
NAMQ	512Holyhe		512 Holyhead Town																	
NAMR	512Kingsla		512 Kingsland																	
NAMS	512Llanba		512 Llanbadrig																	
NAMT	512Llanbe		512 Llanbedrgoch																	
NAMU	512Llandd		512 Llanddyfnan																	
NAMW	512Llaneili		512 Llaneilian																	
Etc	Etc	Etc	Etc																	
Etc	Etc	Etc	Etc																	