

Market Performance Framework

Performance Advisory Group – PAG15

8 February 2023

Agenda – 2hrs (10am –12pm)

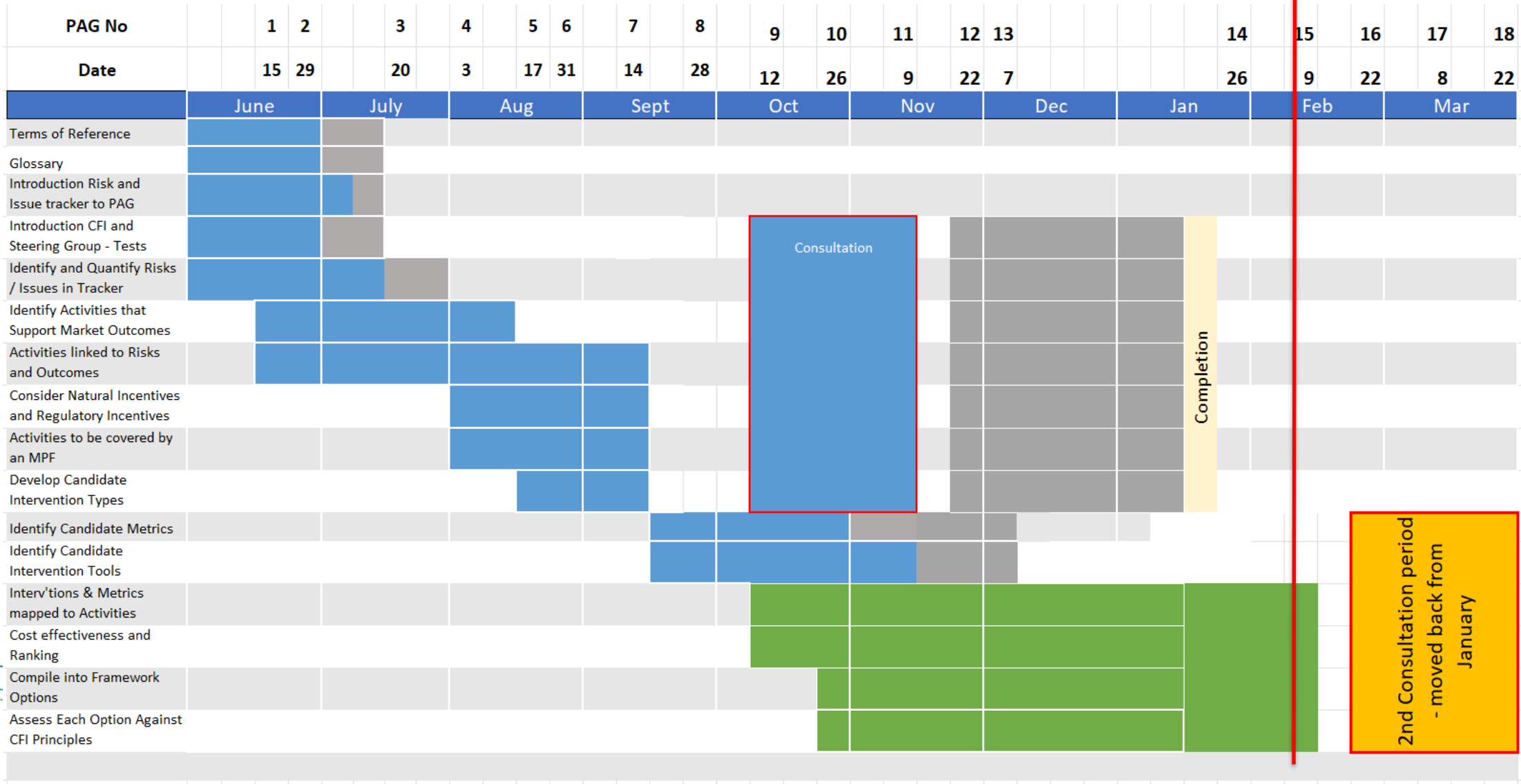
1	Welcome	Chair	5 Mins
2	Actions from previous meeting	Chair	5 mins
3	Review progress to date & Steering Group feedback	Chair	5 mins
4	Walk through of tools descriptions	Axelle Saada/ Janet Judge	50 mins
5	Feedback on Middle tier Activity templates	Axelle Saada/ Janet Judge	20 mins
6	PAG Feedback Metric review	Axelle Saada/ Janet Judge	30 mins
7	AOB	Chair	5 mins

Actions from Previous Meeting (1 of 1) to be updated

NO	Action required	Action by	Action Date
----	-----------------	-----------	-------------

14/01	Feedback received from members on the metrics will be retained for future use and added to Metric review.	AS/JJ	24/02/2023
-------	---	-------	------------

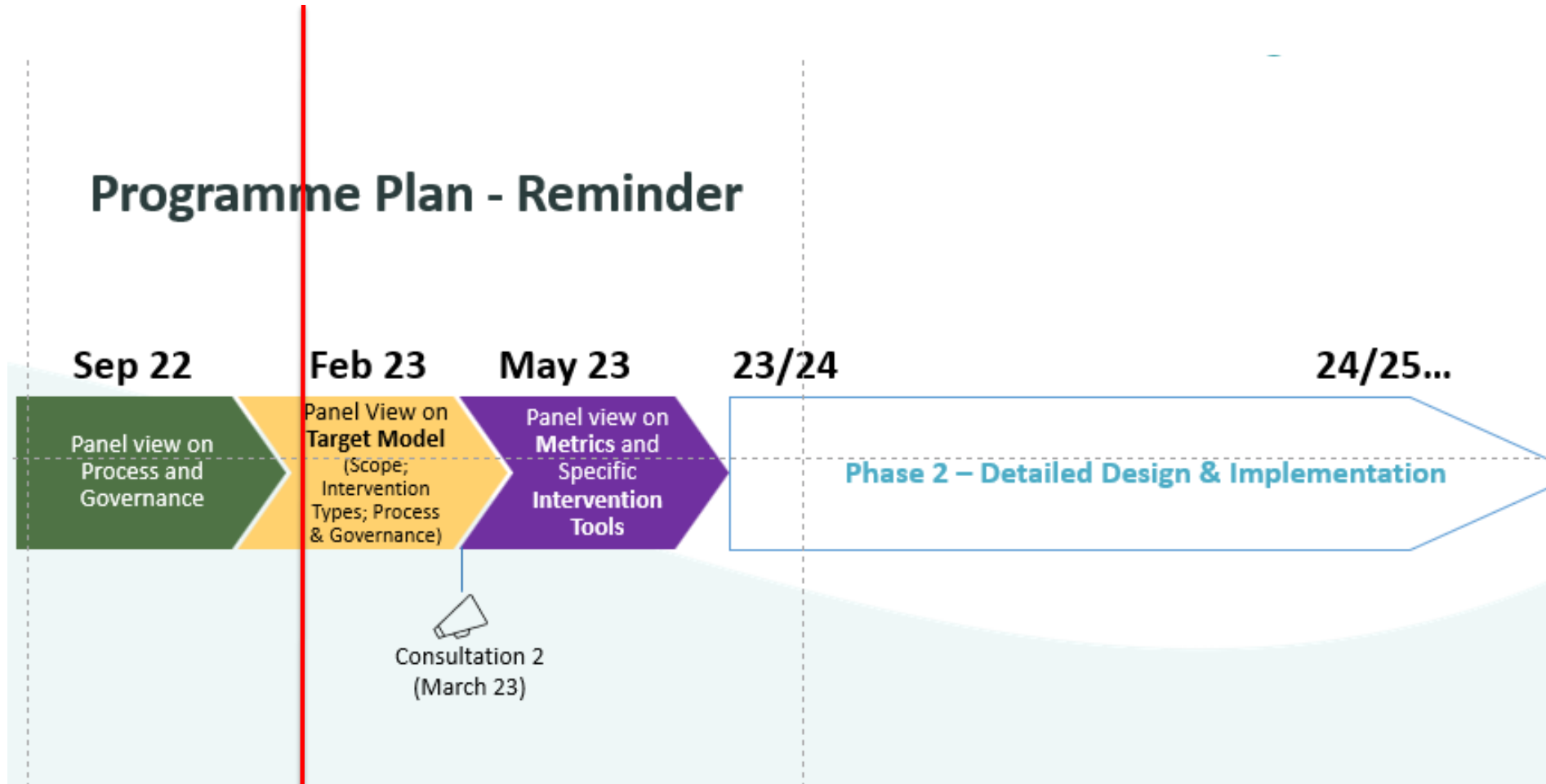
Review Progress to date (1 of 1)



- Not started
- On track
- At risk
- Completed

2nd Consultation period
- moved back from
January

Programme Plan - Reminder



Agenda – 2hrs (10am –12pm)

1	Welcome	Chair	5 Mins
2	Actions from previous meeting	Chair	5 mins
3	Review progress to date & Steering Group feedback	Chair	5 mins
4	Walk through of tools descriptions	Axelle Saada/ Janet Judge	50 mins
5	Feedback on Middle tier Activity templates	Axelle Saada/ Janet Judge	20 mins
6	PAG Feedback Metric review	Axelle Saada/ Janet Judge	30 mins
7	AOB	Chair	5 mins

Walk through of tools descriptions

- For each of the 45 tools there is an overview i.e. a summary similar to the summaries produced for the 11 intervention types
- Each covers both wholesalers and retailers, except where the tool applies to one trading party type only
- The overview templates include:
 - The intervention category and type
 - The name, ID and description of the intervention tool
 - Prerequisites for using the intervention tool
 - Comment on the flexibility with which the intervention tool can be used
 - Details of the cost of the intervention tool based on low, medium, high rating
 - The governance layer(s) applicable to the intervention tool
 - Nuances applicable to the intervention tools
 - Benefits and drawbacks of the intervention tool to stakeholders

Walk through of tools descriptions

- ◆ Seeking endorsement of the tools template
- ◆ The 'ask' of the PAG
 - ◆ Is there anything missing from the template that you believe needs to be included?
 - ◆ Are there any show-stoppers?
 - ◆ Are there any tools missing?
- ◆ We will send more overviews to PAG this week

Walk through of tools descriptions

Examples of tool overviews

- ◆ [Public Peer Comparison \(vs target standard\) OP.docx](#)
- ◆ [Monthly penalty payments OP.docx](#)
- ◆ [Credit Cover Adjustment OP.docx](#)
- ◆ [Limitation of Central System Access OP.docx](#)

Endorses Tools template Endorsement

- 💧 Seek Endorsement – Tools template
- 💧 What Endorse covers – The sections of the template cover what needs to be included i.e. there is nothing missing
- 💧 What this Endorsement doesn't cover – The specific content/wording in the template

Agenda – 2hrs (10am –12pm)

1	Welcome	Chair	5 Mins
2	Actions from previous meeting	Chair	5 mins
3	Review progress to date & Steering Group feedback	Chair	5 mins
4	Walk through of tools descriptions	Axelle Saada/ Janet Judge	50 mins
5	Feedback on Middle tier Activity templates	Axelle Saada/ Janet Judge	20 mins
6	PAG Feedback Metric review	Axelle Saada/ Janet Judge	30 mins
7	AOB	Chair	5 mins

Feedback on Middle tier Activity templates

- Feedback on Middle tier Activity Templates?
[Middle Tier](#)

REF	Activities and Behaviours
CV.13	Market participants improve processes to identify gaps sites and bring into charge
CV.16	Market participants work collaboratively to ensure vacancy classification is accurate and resolve associated issues
CV.18	Provision of timely and accurate tariffs into the central systems
GS.10	Market participants to support customers during unplanned events, including appropriate maintenance and use of customer emergency contact details
CV.1	Retailers to handover good quality timely customer data and incoming retailer to take a timely actual transfer reads on customer switching
GS.5	Wholesalers better reflect the needs and requirements of their non-household customers in the provision of non-household services, including trade effluent and planned/unplanned events
GS.8	Wholesalers to provide timely, clear and concise explanations around most common customer wholesale-related queries
CV.7	Retailers to report customer queries to wholesalers in a comprehensive and timely manner
CV.8	Retailers to keep customers informed on query resolution
CV.23	In the absence of a meter read retailers should make sure an accurate estimation of consumption is made
GS.2	Retailers provide clear information to customers on how the complaints process works
GS.6	Wholesalers provide clear and accessible information on their relative performance on non-household services

Agenda – 2hrs (10am –12pm)

1	Welcome	Chair	5 Mins
2	Actions from previous meeting	Chair	5 mins
3	Review progress to date & Steering Group feedback	Chair	5 mins
4	Walk through of tools descriptions	Axelle Saada/ Janet Judge	50 mins
5	Feedback on Middle tier Activity templates	Axelle Saada/ Janet Judge	20 mins
6	PAG Feedback Metric review	Axelle Saada/ Janet Judge	30 mins
7	AOB	Chair	5 mins

Feedback on Metrics

- You gave us over 120 comments
- Provided interesting insight, adding value by helping us improve metrics
- Please make sure that your feedback is argued so we can best act on it (e.g. if you don't like some aspect of a metric, tell us why and how to improve it)
- A number of comments have been put on hold and will come into play at a later stage

Feedback on Metrics

- CV.2 / UWE.2: Retailers to make sure that CMOS is regularly updated with good quality customer & consumption data [Assumes meter reads have been completed]

Name	What is being measured	Comment	MOSL response
Number of vacants with consumption		Yes, but would need to have consideration for any vacant with consumption SPIDs that have an outstanding bilateral raised for them.	We can remove SPIDs with an outstanding bilateral and with internal meter to get a cleaner metric. Combined metrics will be discussed in the next stage.
		Although vacant with consumption indicates an issue, the retailer on its own cannot always rectify the situation, eg when premises are vacant and there is a leak. This could lead to Retailers being unfairly penalised because they can't rectify the situation – what mechanism is there to remove these? Does focusing primarily on vacant with consumption create the perverse incentive for retailers not to read meters during vacancy? Can vacant premises with internal meters be excluded because the meters are impossible to read?	We also monitor volume of reads for meters at vacant premises to ensure we understand performance in this area and that this is being maintained. Will add as a new metric to be looked at alongside this one.
		I would also focus on meter reading during vacant periods. Current MPF stops the clock on KPI days elapsed when vacant. If it hasn't been read, how do we know there is consumption?	
Number reads removed after subsequent read inputted		No. Does not give a meaningful indication of quality. Furthermore, removal of reads after a subsequent read is inputted may be a sign that trading parties have corrected a long unread or unreadable meter. This is a good thing and something we do not want to discourage.	This is only proposed to be used to trigger audit or additional monitoring. We will look at all potential perverse incentives of metrics in the next stage.
Proportion of rejected reads	Proportion of rejected reads (nuanced by rejected reads for SPIDs that had Volumetric Adjustment, occupancy status change or customer name change)	Unsure what this metric is meant to be representing. What does a rejected read have to do with customer name change or occupancy status?	All these factors could cause the read to be different than what CMOS expects and therefore cause rejection even though the read is correct.
Number of reads resubmitted	Number of times new read submitted with same date and new value where a read was previously rejected	Need to be careful that we do not discourage corrections	This would be a positive metric (exactly to encourage corrections)
		Times new reading submitted with the same date and read is dependent on CMOS rejection criteria which is not accurate and potentially puts MOSL and TP in conflict. Criteria must not be subject to doubt or counter claim. Remove.	We recognise that a read can be correct but fail validation. What we're looking for here is the number of times new read submitted with same date and a new value indicating that the read originally processed was investigated then corrected

Feedback on Metrics

💧 CV.2 / UWE.2: Retailers to make sure that CMOS is regularly updated with good quality customer & consumption data [Assumes meter reads have been completed]

AS/JJ
11:55

Name	What is being measured	Comment	MOSL response
LUMs		What is the baselines for good/bad performance. Will it take into account active bilaterals? External vs external needing split. Must be viewed against similar performance by the wholesaler, e.g., address accuracy, meter location or details etc.	All these nuances will be fleshed out in phase 2
		I would still challenge our calculation of LUMs and I think there are arguments that we may in certain areas overstate the problem. I support LUMs as a metric but the underlying calc I would question.	
Settlement accuracy	Settlement accuracy	Is this a double whammy with YVEs and MPS 15,18,19?	Propose to remove this as a secondary metric for underperformance payments. This would be a trigger for recification, where performance is deteriorating.
		Yes, but more detail needed. What would the baselines be, don't want to incentivise unwillingness to correct incorrect information.	
YVEs	Number of YVEs in the system and how often are they updated	Depends, what else is considered. Minimum YVE updates per year, move ins, leaks, vacant with consumption?	YVEs are relevant as this activity is about retailers entering and maintaining data on CMOS. All nuances will be fleshed out in phase 2
		Even if YVEs are being updated regularly, is this necessarily meaning that customer has access to accurate consumption data?	
		YVE accuracy. Benchmark against historic and/or when actual volume is captured through a read.	
SEMD flag update when new customer moves in		what does this tell us? How does MOSL know if it is correct?	Is marked as a market indicator. Downgrade suitability to 2. Used as trigger for rectification and audit. It would be done by comparing retailers - e.g. if one retailer NEVER updates SEMD flag when new customer name whilst other have a 10% update rate. We could refine this analysis by SIC codes (some SIC codes more likely to have a SEMD flag)
		this flag does not always need updating when a new customer moves in. How would this metric indicate whether or not a SEMD flag update was needed or not?	
CCW customer complaints data	Volume of complaints within subcategories "Incorrect account information", "Estimated billing" and "Disputed meter reading"	Must distinguish cases raised to CCW that relate to inaccurate customer and consumption data being the route cause.	We have specified we would use volume of complaints within subcategories "Incorrect account information", "Estimated billing" and "Disputed meter reading". This is likely to be reported in peer comparison as complaints per 10,000 supply points, as we currently do for holistic reporting. This would be a trigger for audit and or recification. It does not necessarily need to be MOSL data and does not matter whether it is used in other places. If a TP is performing badly audit and or rectification may be appropriate intervention
		Maybe. What complaints data would be used? If just number of complaints then this is too simplistic. Reason for complaint, time taken to resolve, whether one or more trading parties were cause of the issue.	
		If customers are experiencing errors with account (e.g; wrong or incomplete address) then it's an issue first and foremost with retailers systems, but also likely to indicate this is being mirrored in CMOS.	
		CCW complaints data is already using in R-Mex and other reports and not in the control of MOSL or MPC. Recommend remove	

Feedback on Metrics

- CV.2 / UWE.2: Retailers to make sure that CMOS is regularly updated with good quality customer & consumption data [Assumes meter reads have been completed]

Name	What is being measured	Comment	MOSL response
Number of meters that have bilateral tasks raised against them		Would like clarity on what sort of bilateral tasks being referred to – If the activity assumes meter reads have been taken, why has a bilateral task been raised?	Indeed - remove from CV.2 but keep in CV.6
Average timeframe between read rejected and re-read submitted		I fundamentally reject the assumption made here. Not all rejects need investigating. Yes, but only if the metric differentiates between rejection codes and therefore have different baselines for performance for the different rejection codes?	This is based on the average time taken to resubmit a re-read. Suggest we would look at the aggregate but could use breakdown by rejection codes for education and/or rectification purposes or as an additional metric for diagnostic purposes
Time to get completeness of retailer-owned data after gap site/new connection	Average time taken by Retailer to input customer name, SIC code etc after gap site/new connection created (Data Quality APIs)	This is heavily caveated by the accuracy of the underlying CMOS and how early the Wholesaler creates a SPID How would this be measured. What metric would be used to confirm if the customer name was needing to be updated, or alternatively if the gap site/new connection was legitimately vacant and meant to be 'No Customer'.	Remove from CV.2. Instead, add as a market indicator to CV.13 and CV.15
Date of last data item update for retailer-owned items		this is too simple as a measure of quality. What would the baseline be and would you have a baseline for individual transactions? How does frequency or date of last update by retailer indicate quality of that update?	This wouldn't be looked at at SPID level but look at the distribution of last data item updated date for all SPIDs and compared across retailers. Quality of the update will be measured by other metrics in the list.
Volume of CMOS updates following completion of a bilateral process		No, what would the baseline for this metric be? Is more updates after a bilateral completion good or bad? Is there always an expectation of CMOS updates after the completion of a bilateral, and if yes, would those updates not be more likely to come from a wholesaler?	Amend suitability and cost RAG to 1 (red). Leave in list of metrics but remove from interventions [audit or an additional metric monitoring intervention.]
Number of backdated bills	Number of backdated bills beyond X number of months that had to be issued to customers	This is a CMOS related activity what do customer bills have to do with this? No, unless this metric included the route cause issue as to why the backdated bill was needed, and which trading party was responsible.	Propose to move this metric to UWE.4 and use as trigger for audit and rectification.
USRs	Volume by TP pairing	The USR may be as a result of wholesaler activity	Added party who pays for USR and party who benefits from it as additional factors to record.

Feedback on Metrics

- CV.3 / UWE.3: Wholesaler to make sure that CMOS is regularly updated with good quality accurate asset and premises data including timely deregistration of premises not eligible where appropriate

 AS/JJ
 11:55

Name	What is being measured	Comment	MOSL response
General comment		Should meter tampering also be considered as a metric to be created?	How would you measure meter tampering? Don't think this is measurable
		An additional metric that might apply to all interventions could be retailer self reported billing or administration complaints (relating specifically to this area) that retailers have attributed to wholesalers. Retailers currently report to CCW which complaints are wholesaler attributable, but they may be able to break this down further.	Add "self-reported retailer complaints on billing or administration that retailers have attributed to wholesalers.
UPRN and VOA	Proportion of UPRN and VOA codes (inputted and validated)	Not sure I understand the 2 rating for cost of the metric	Cost of 2 - due to validation process requiring MOSL or third-party analysis
		Yes, presuming someone like Sagicity are being used to check the data. Also need to view the number of UPRN updates that change the UPRN completely to a new address, rather than just a typo correct.	We can add proportion of UPRN that were changed to a new address as a sub-metric, but as is proposed for financial incentives we need to ensure that we do not discourage wholesalers from making UPRN corrections
GIS coordinates	GIS coordinates for meters (inputted and validated)	Yes, but how will the accuracy of this data item be measured?	Based on analysis from MOSL or a third-party (e.g. matching GIS with address). May include checking against retailers' data records.
Unpaired SPIDs		Not sure I understand why suitability is only a 2	Difficult to determine whether a SPID should be paired or not. Propose to exclude D2086 where code is NE_NOSEWERAGE and NE_NOWATER.
Address of premises	Proportion of addresses (inputted and validated)	Not sure I understand the 2 rating for cost of the metric	One of the key areas of central cleanse is to provide the MPF a market approved data quality scorecard that we can measure TP owned data item quality.
		Need to decide on what is central data cleanse (and managed through that programme) and what is MPF. It might be that these metrics need to be on hold until that project is complete. Too many driving forces that have a risk of a clash.	
		Yes, presuming someone like Sagicity are being used to check the data	
Legacy LUMs		Yes, but would need to be combined with the data sets reviewing validity of premise and asset data on the SPID.	This activity is on wholesalers and premises data is wholesaler-owned so these additional metrics would not shift accountability.
R-MeX	Scores on Quality of data maintenance and improvement category	Yes, although very subjective as a metric.	Not proposed to use for financial incentives or sanctions
Late meter reads and missed reads	MPS 7 to 14	Yes, but would need to be combined with other data sets to present clear analysis on the cause of the read being late or missed.	This wouldn't trigger interventions on a SPID level. We would look at the number of reads that get missed or delayed and look to rectify if poor performance.

Feedback on Metrics

- CV.3 / UWE.3: Wholesaler to make sure that CMOS is regularly updated with good quality accurate asset and premises data including timely deregistration of premises not eligible where appropriate

AS/JJ
11:55

Name	What is being measured	Comment	MOSL response
Volume of meter issue rectifications (OPS Bs and Cs)		Data item capture from completion transactions. Not all B's and C's. Faulty meter replacements may creep into overlaps with capital replacement etc.	Remove this metric and amend metric on volume of CMOS updates to include more B and C processes.
USRs	Volume by TP pairing	Why is suitability for this a 2 but for retailers a 3?	Because incorrect reads are the main cause of USRs and retailers are responsible for taking reads. Wholesalers have a responsibility in facilitating those reads.
		This could also be driven by retailer data quality. We could assume by who pays for it as the responsible party but not sure we record this.	Added party who pays for USR and party who benefits from it as additional factors to record.
Volume of CMOS updates following completion of a C1 (and other relevant processes such as C5)	Volume of updates but also timeliness, completeness and accuracy if possible to report on	What is the baseline for this metric? Is more updates after a C1 bilateral completion good or bad? Could be useful but could also encourage behaviour that may be detrimental.	This would be compared across wholesalers and used to trigger an audit or an additional metric monitoring intervention.
		Completeness and time to update should be in addition to volume of CMOS updates.	Added timeliness, completeness and accuracy as additional factors to record.
NaPs report	SPIDs that sit in the New and Partial pool	No, this would not really indicate whether or not the SPID is sat at 'New' or 'Partial' because of address or asset issues. Incorrect address and asset data would also not stop a SPID from becoming tradable, but would more impact the ability for the retailer to move in a customer/verify the customer, and collect meter reads.	The MIF project on NAPs highlights joint issues. It is predominantly pointed at wholesalers where the address and premises data quality is poor and cannot be matched to external datasets (circa 73% of NAPS. While a retailer could approve a SPID with poor data, it doesn't stop us calling this out ongoing within the NAPS
Date of last data item update for wholesaler-owned transactions		Why is suitability a 3 for wholesalers but a 4 for retailers?	Because retailer data items get updated more often.
		This metric by itself would not indicate whether or not indicate whether there asset and premise data are correct or if the wholesaler is performing well	This wouldn't be looked at at SPID level but look at the distribution of last data item updated date for all SPIDs and compared across wholesalers.
Deregistrations in CMOS that are not matched a C5 ORID		[...] A dereg can off the back of other processes and it is a definite for permanent I processes. There is sense in terms of efficiency if a dereg is identified off the back of a C1 and does not affect another wholesaler or retailer, that the dereg can be proposed and agreed without the need for a C5.	Propose to remove metric. Retailers have mechanism e.g. disputes by which they can challenge if wholesaler deregistering SPIDs without their knowledge. MPF will focus on SPIDs that should be deregistered
		Please clarify the specification for the deregistration metric. Is this measuring the timely processing of all de-registration requests? Given the large part that retailers also play in market eligibility requests, has this been taken into account, given the impact that a retailer decision/action can have on these being processed.	Will use proportion of UPRN and VOA as primary metric for under-performance payments.
Proportion of assured LTV		This would need to be combined with other metrics such as address, UPRN and VOA validity and accuracy reporting to indicate whether the long term vacancy is being caused by incorrect premise and asset data	Amended the metric to proportion of assured LTV which will enable the wholesaler to put in the verification work

Feedback on Metrics

- CV.6: Retailers to provide sufficient level of actual meter reads to ensure accurate consumption data (giving accurate settlement and customer billing)

AS/JJ
11:55

Name	What is being measured	Comment	MOSL response
General comment		Multi-site customers – disbenefits accrue where any subset of meter reads for a multi-site customer are late and/or inaccurate. Do / can the metrics or interventions capture this? Customer own reads - Assume accuracy and timeliness is applied to these.	It will be difficult to measure meter readings by customer as we don't have this information in CMOS. Accuracy and timeliness applies to customer reads in same way as any other read methods.
Proportion of estimated transfer reads		No. What would the baseline for this be? Would need to include any issues with gaining a visual read due to wholesaler outstanding work or bilaterals, which itself will have widescale geographical variation from Wholesaler to wholesaler.	Amend metric to exclude supply points where there is a relevant open bilateral request
Proportion of estimates in R2, R3 or RF		Yes, but like the MPS 15 to 19 metric, data should include where that SPID is awaiting a bilateral or other rationale that would explain why the read might be an estimate. May well be outside of the retailers control.	This as a market indicator. We could do this on an aggregated basis e.g. proportion of estimates and volume of open bilat requests. Accounting for this on SPID level (i.e. this SPID was settled on estimate and it has an ORID therefore we remove it from the count) may be overkill. We can add the option into the metric and decide based on cost and benefit later.
Volume of transaction corrections in relation to meter reads		No, what would be the baseline. Lots of corrective transactions shows an intent to correct bad information in the market. Many cases of bilaterals result in wholesalers requesting removal of reads to remove/replace incorrect meters. Would the retailer be penalised for that. To simplistic to be a metric.	We are using this as a market indicator. We will remove from under-performance payments
LUMs		Yes, but must include the reasoning why a metering is long unread, and which trading party the issue is sat with, e.g. outstanding or deferred bilateral?	All these nuances will be fleshed out in phase 2
Billing complaints relating to consumption	Both CCW (Estimated Billing, Disputed Meter Reading, Disputed Liability and Meter Reading Frequency) and self-reported	Yes but this must include the reasoning for the complaint. Many complaints raised relating to consumption can be because of the tens of thousands of long unread meters being resolved and causing significant impacts to bills once visual meter reads are submitted and used.	We have LUM/LLUMs as a metric in the list. Combining metrics will be looked at in the next stage.
Skip codes data		MOSL may have underestimated the cost of this metric. To analyse the skip codes would be expensive	Amend RAG to 2 (amber) as we don't have skip code data, would be manual task to get data from retailers and analyse but cost would be reduced if retailers use the standard skip code list and process
USRs	Volume for each retailer	Yes, but would need route cause analysis combined to determine which trading party was responsible for the need for a USR.	Added party who pays for USR and party who benefits from it as additional factors to record.
Meters that have bilateral tasks raised against them		Yes, including breached SLAs and deferred bilaterals.	We will count meters with open bilateral task so that we can remove these from other relevant metrics in the list of metrics for this activity

Feedback on Metrics

💧 CV.9: Wholesaler to resolve request in timely and efficient manner (with regards to customer needs) to the appropriate level of quality

AS/JJ
11:55

Name	What is being measured	Comment	MOSL response
General comment		Will metrics be superseded by BR-MeX?	Metrics that overlap with BR-MeX will be removed from any financial incentives but can still be considered for other incentives. Rejections and deferrals will be looked at on an aggregated basis, not by ORID level to allow for an appropriate of rejections/deferrals to take place.
		Overall the key for this risk is around timeliness e.g. overall bilats performance & quality e.g. BR-MeX.	
		Need to be mindful of using rejections/ deferrals etc as these may be rightly set to this status.	
Customer enquiries (F4 process) bilats data	Volume and timeliness of progression through process	This should be speed of substantive response not the volume of enquiries received.	Several elements of the bilats data will be taken into account indeed (e.g. volume and timeliness of progression through process). We propose to downgrade the suitability of F4 from 4 to 2.
		Potentially. Are customer enquiries an indication of good or bad service. Not necessarily, will details of enquiry and how it was resolved be reported on in this metric.	
Customer complaints (F5 process) bilats data	Volume and timeliness of progression through process	Yes, but dependant on what data is used from the customer complaint.	Volume and timeliness of progression through process
Meter and supply arrangement verifications (C1 process) bilats data	Including timescales for responding and data on deferred bilat requests but also number of Retailers resubmitting requests and disagreeing with proposed outcomes	Needs further definition. There may not be a problem.	Added detail to what is being measured - Including timescales for responding and data on deferred bilat requests but also number of Retailers resubmitting requests and disagreeing with proposed outcomes. Using as a trigger for rectification. We can look at weightings in phase 2
Volume of customer contacting wholesaler and number of wholesaler-raised F4 and F5		No. Volume of customer enquires and complaints raised by itself meaningless. Would need inclusion of why the F4 and F5 were raised.	There are definitions of when a wholesaler should raise and need to be careful we don't discourage or penalise for raising. Amend metric to be volume of F4 and F5 s this is an indication of an underlying issued. Could include category from bilaterals.
Volume of CMOS updates following completion of a bilateral process		There must be a shortcoming in the design of the Bilat hub if this is a high cost. This metric should also be about time taken to update CMOS too	Cost is high because both systems don't talk to each other.
		No. How do you determine and baseline how many CMOS updates are needed after completion of a bilateral? May incentivise trading parties to reduce or limit number of updates after a bilateral.	We would look at this on an aggregated basis, not by ORID

Feedback on Metrics

💧 CV.9: Wholesaler to resolve request in timely and efficient manner (with regards to customer needs) to the appropriate level of quality

AS/JJ
11:55

Name	What is being measured	Comment	MOSL response
Disputes data		Yes, dependant on what data from any disputes is used, e.g. outcome, reason for Dispute, length of time taken to resolve etc.	Will add these elements to the metrics definition
Wholesale policies regarding allowances to customers for time spent with issue		I think this needs looking at elsewhere. I think the main point for the MPF is the discussion that is needed on redistribution of market charges as this has always come up as an option. Its not a metric.	This one was suggested by Wendy in the previous round of feedback on metrics. Mapped to compensatory payments. Discuss with PAG.
R-MeX	Scores on speed and quality of responses to requests	RMEX could change to BR-MEX under PR24, which could be a possible metric across all the interventions. Depending on the eventual design, this may provide direct customer insight into how well wholesalers are resolving requests.	Once BR-MeX is further defined, we can look at overlap with proposed metrics and update accordingly.
Number of rejections that happened on the last day of the SLA		It's a valid thought but when you set a KPI, you set that it is within the right of the TP to use that time as they see fit. Its not best practice and creates risk obviously to leave things to the last minute. I would focus this on the accept/reject time efficiency as I have suggested before that for any bilats you should review within XX days of receipt. This creates a pace in a request and allows for keeping the customer informed.	Wholesaler rejecting a request on the last day of the SLA is an indication of poor behaviour. We could look to track the number of rejections subsequently rejected by retailers and the reasons for this (using D8231)
Average number of deferrals per ORID	May need to be looked at by TP pairings	This raises the point that overall time delivery is a shared responsibility. Deferrals can be due to the other trading party and some deferrals are unavoidable to a point.	We may need to refine the metric to look at TP pairings but this can be elaborated further in the next stage. We would look at ORIDs on an aggregated basis.
Proportion of ORIDs that get resubmitted (after rejection, and after completion)	May need to be looked at by TP pairings	You cannot assume that a resubmission is a negative on the wholesaler. The wholesaler may have rejected due to missing info that the retailer subsequently provides as part of the resubmission. Need to look at deferrals and rejections as a shared metric between pairings. Its a collaboration piece.	We may need to refine the metric to look at TP pairings but this can be elaborated further in the next stage. We would look at ORIDs on an aggregated basis.
CCW complaints data	Volume by wholesaler region, reason for complaint, time taken to resolve, whether one or more trading parties were cause of the issue	Maybe. What complaints data would be used? If just number of complaints then this is too simplistic. Reason for complaint, time taken to resolve, whether one or more trading parties were cause of the issue.	Will add these elements to the metrics definition
Number of days in addition to the "happy path" SLA		Yes, but who determines the 'happy path' SLA. Is this the code defined?	It is indeed referring to the code SLAs; the happy path SLA is the overall number of days for a process if no SLA is failed

Feedback on Metrics

- CV.15: Market participants improve processes to facilitate new connections (including non-eligible customers post development or change of use)

Name	What is being measured	Comment	MOSL response
General comment		Not much here on exits. Again possible double up on central data cleanse.	Do PAG have any suggestions on relevant metrics? Would no. leakage notices acted on by retailers work?
		Is there a way of measuring how many de-registrations occur as a result of a premises being left incorrectly in the market post new connection? Thjs may be another useful indicator of customer impact.	Do PAG have any suggestions on relevant metrics?
Vacancy performance on new connections	How long new connections remain vacant and how frequently that happens	What metric is this? How would a baseline of how many vacant new connections numbers be determined? Combine this metric with consumption data to show vacant sites that have consumption.	This is only proposed to be used to trigger audit. This metric is a market indicator. We propose to use market indicators as triggers for discretionary interventions for intervention types of education, additional metric monitoring and audit
Completeness and time to input retailer-owned data	Average time taken by Retailer to input customer name, SIC code etc. (Data Quality APIs)	No. This metric is too simplistic. Other data would be needed to combine into this metric to show whether or not it is the fault of the trading party, or (highlighted in the risks for this activity) whether it might be the customer hindering the process, or whether the SPID is correctly marked as vacant.	We propose to monitor wholesaler and retailer data as TP pairing
NaPs performance	Amount of time from new to partial and from partial to tradeable - refine by overall volume of new connections	Yes, but compared to market defined SLAs for this process and would need to highlight and separate delays caused by either the wholesaler or retailer.	Monitor metric as TP pairing
R-MeX	Scores on Quality of data maintenance and improvement + Level of engagement and support	Recommend being mindful of using R-MeX - can be a subjective measure	Not proposed to use for financial incentives or sanctions
Unpaired SPIDs	Volume of unpaired SPIDs at point of connection and proportion that later got paired	Need to work through complexities as to who is the target for the measure for cross border supplies.	Reduction of unpaired SPIDs is good for the market. Cross borders supplies are a small proportion. Used as a trigger for rectification only.

Feedback on Metrics

💧 GS.3: Retailers resolve complaints in a timely and efficient way

Name	What is being measured	Comments	MOSL response
General		Adoption of [RWG] customer complaints GPG [Good practice guidance] and complaints MoU [Memorandum of understanding]?	Add to metrics. Suggest map as a secondary metric to publishing peer comparison
Difference between volume of retailer complaints and volume of complaints escalated to CCW		Yes. The data already exists and in part if provided to CCW, however this would need a baseline for performance defined, and with a basis taken from the Customer Protection Code of Practice.	Baseline performance required where performance is compared to a target or for financial incentives. This will be defined when we develop tools and metrics further
Average "age" of complaints		This metric by itself does not paint a good enough picture to determine if a complaint has been handled well or not. The route cause of the complaint needs to be included as well as quality and effectiveness of how the complaint was handled.	Can do some segmentation of complaints now based on CCW categories. May be able to obtain this at a more granular level if we were to ask CCW (agents) to amend their system
Self-reported retailer complaints data	Data on timescales for different complaints milestones	<ul style="list-style-type: none"> • Yes, but the focus should be using effectively using the data already supplied by retailers to CCW. In particular the timescales for each stage of a complaint, i.e. stages 1-4. The focus must be on the Customer Protection Code of Practice baselines for performance, not arbitrary new performance standards placed on top of these. • Often it is the customer or the wholesaler that can prolong resolution • Key to meaningful and transparent performance data is consistency, what will ensure Retailers are consistently identifying contacts as complaints? 	<p>It is the data already reported by retailers to CCW that we're proposing to use, which we'd get from CCW. Everyone is prone to the same delays. Can factor in CCW wholesaler and retailer classification.</p> <p>There is CCW guidance on complaints that is regularly updated and good practice from RWG on complaints that help bring consistency.</p> <p>Detailed design of metrics and defining standards will be part of phase 2.</p>

Feedback on Metrics

💧 GS.3: Retailers resolve complaints in a timely and efficient way

Name	What is being measured	Comments	MOSL response
Volume of GSS failures regarding regulations 17C and 17D	Both how much did the retailer pay out automatically and how much did it have to get chased for	This is data that has never been requested from retailers since market opening. Unsure how this metric would support the intervention and presenting a clear picture as to the quality and effectiveness of how the complaint would be handled.	Provides additional insight on performance on timeliness
F5 bilateral data	How quickly retailers and wholesalers are resolving complaints, e.g. how long before the retailer closes the F5, how quickly the retailer responds to a wholesaler action in the bilateral hub, etc.	<ul style="list-style-type: none"> • Yes, could do with measuring wholesaler effectiveness. What data would be used as part of this metric? F5 bilaterals more a focus on the wholesaler than retailer, but would need to include data to determine if the issues is with the retailer, wholesaler, or both. • Inclusion of F5s needs to be considered carefully as either a performance metric or a market one. 	<p>Details of the metric already provided e.g. duration taken by the retailer to close the request.</p> <p>This is an easy to track performance metric.</p> <p>We could also consider tracking 'Info request' and how quickly the retailer is responding.</p>
Customer survey on complaint resolution		Yes, but this is a very subjective and geographically impacted metric.	The survey would be designed in order to give robust results, including across regions.

Feedback on Metrics

- GS.9: Wholesalers efficiently maintain working/functioning assets, including meters which must be working accurately and be readable

Name	What is being measured	Comment	MOSL response
B5s bilats data	Number of B5s raised for that wholesaler and timeframe of resolution (could include other B and C processes like installation and accuracy testing)	Why is the cost a 3, doesn't the bi lat hub produce this data already?	This data exists in the hub but the reports would need to be created
		Must include number of deferrals as well.	Add number of deferrals to metric
F4 (enquiries) and F5 (complaints) bilats data	Number of F4s and F5s raised for that wholesaler and timeframe of resolution	Inclusion of F4/F5s needs to be considered carefully as they can include complex queries/ complaints that are not easily categorised.	Change metric suitability to 2
		Yes, but needs granular details such as outcome of F4 or F5, and the time taken to rectify issue.	Add suggested granular details to metric description
USRs and settlement accuracy	Including % SPIDS settled on actual reads at R1, R2, R3 and RF	Feel as though metrics around settlement accuracy including actual reads has not been included in the interventions. Surely this is a key indicator of issues with assets.	What intervention types would you link that to?
Volume of GSS failure on reg 17G, 17H and 17I (low pressure and sewer flooding)		Need to ensure we are not measuring things twice	We wouldn't be re-measuring GSS failures, only gathering the existing measures to assess performance
Number of occupied premises with zero consumption		No, this is too simplistic. Would need to have other data included, such as size or type of premise, physical assets that use water within the premise confirmed. How would temporary holiday homes be factored. Premises that may have wide ranging consumption use changes, many of which may have 0 consumption for periods of time.	Change metric suitability to 2 and remove it as relevant metric for financial and reputational incentives
		Provided it has been read? Also it might be incorrectly occupied.	Agreed. This could be one of the factors the audit looks like to determine accountability and next steps

Feedback on Metrics

- UWE.4: Retailers facilitate the provision of timely and accurate consumption data to customers and support them in resolving supply pipe leaks [assumes meter reads have been completed]

Name	What is being measured	Comments	MOSL response
General comment		No metrics appear to relate to resolving supply pipe leaks	Do PAG have suggestions on relevant metrics for measuring this?
Retailer audit	Evidence of consumption data on customer communications (including bills)	Cannot comment as no detail provided. 'Evidence of consumption data on communication with customer'. What would the minimum standard be? Who would measure these standards?	Governance will be part of consultation 2. Setting of standards would be part of phase 2
CCW complaints	Volume of complaints in categories Disputed meter reading, Estimated billing, Leak allowance and Meter read frequency	<p>CCW complaints need to be specific to consumption data.</p> <p>No, if the only focus is on the volume of complaints raised. Is this an indication of whether the complaint was legitimate, and whether the trading party at fault was clearly indicated.</p> <p>This data will be provided to MOSL by CCW and much is already provided so not sure why this is a 2 and not a 4 for cost.</p>	<p>Metric would be volume of complaints per 10,000 for categories Disputed meter reading, Estimated billing, Leak allowance and Meter read frequency</p> <p>Everyone is subject to the same process. Can ask CCW whether they do/can record whether the complaint is related to wholesaler or retailer, in same/similar way to retailer self-reported complaints</p> <p>Cost of metric is 2 to allow for this extra analysis work to cover aspects that aren't currently collected/reported on.</p>
Vacancy reporting		No, should this not be vacant with consumption? What does vacancy have to do with facilitating the provision of timely and accurate consumption data to customers? If a SPID is vacant then there is no customer.	Agree with suggestion. Amend to vacant premises with consumption as these are potential customers that aren't receiving a bill or premises where leakage is occurring

Feedback on Metrics

- UWE.4: Retailers facilitate the provision of timely and accurate consumption data to customers and support them in resolving supply pipe leaks [assumes meter reads have been completed]

Name	What is being measured	Comments	MOSL response
Completeness of customer name/banner name (or other elements like SIC code)		<p>This metric would be more suitable for CV2</p> <p>Yes, but current way customer name completeness is measured in Holistic reporting is flawed. Metric and data used for this should be reviewed.</p>	<p>Is used in CV.2. Is also suitable for UWE.4 as it helps to correctly identify the customer which is important for understanding water consumption/efficiency.</p> <p>What changes to customer name completeness would PAG suggest to improve the measure?</p>
Measure of retrospective accuracy of meter reads and estimates	Including meter reads removed and backdated YVEs	<p>I don't think that this is a 3 rated suitability it is a 2 against this activity. My thinking is that this is about info to the customer not the data in CMOS</p> <p>How would this be measured and with what data? Need to drive the correct behaviour, and punishing trading parties that may need to alter historical reads/estimates could be detrimental.</p>	<p>Have retained as a 3. Whats in CMOS and it's accuracy is relevant to understanding water consumption and helping to identify leaks, particularly if a customer switches.</p> <p>There are no financial interventions associated with this metric. It has a reputational intervention of peer comparison. The specific tool is still to be determined, and would be used as a trigger for audit and assurance.</p> <p>Once we've defined the specific tool and further defined the analysis that will form the metric, we can decide it's not robust enough for peer comparison but at this stage we're minded to leave this in.</p> <p>The design of the metrics will be covered in phase 2 of the programme where we will consider accountability, dependencies and unintended consequences.</p>
USRs and settlement accuracy	Including % actual reads at R1, R3 and RF)	Yes but can this metric include data that fairly indicates which trading party may have caused any USRs or settlement inaccuracy.	<p>Amend to market indicator.</p> <p>Here we are using only as a secondary trigger for audit.</p> <p>For USRs add: The party who paid for the USR and the party who benefitted from it can be indicators of responsibility.</p> <p>If/where we can determine the cause of parity issues we may be able to determine which TP is the cause. We can add this option into the metric and decide based on cost and benefit later.</p>

Feedback on Metrics

- UWE.5: Market participants work collaboratively to share and utilise additional data (e.g. wholesaler meter reads, SIC) to ensure accurate (consumption) data is made available to facilitate efficiency offerings to customers

Name	What is being measured	Comments	MOSL response
General comment		I would suggest a number of these metrics are covered elsewhere. If Rmex is effective, it should give a view of collaboration on data sharing etc but it is only a 1 way mechanism/capture of feedback. I think this area should be about gaps not covered elsewhere and that add value EG Contact info for unplanned events etc. Apologies, just think meters and data are getting enough attention elsewhere	Metrics duplication across activities is not a problem at this stage. There will be an exercise of metrics rationalisation at the next stage.
		There needs to be an additional metric (potentially applying to all interventions) that clearly measures the water efficiency information being provided to customers, as this is the clearest indication of how well trading parties are collaborating in this space. Perhaps an audit could uncover what type of advice is being provided to customers either through 'self-serve' (e.g; website) or direct contact.	Add retailer and wholesaler audit as a metric
Wholesaler reads volumes and rate of adoption by retailers	By TP pairing	Yes, but should include quality checks of W read if not used by retailer.	Can be controlled to some extent by looking at this by TP pairing. We would look at this on an aggregated basis.
Completeness of customer name/banner name (or other elements) + time between the effective from date and the transaction date		No. This metric has nothing to do with the activity. Market participants work collaboratively to share and utilise additional data (e.g. wholesaler meter reads, SIC) to ensure accurate (consumption) data is made available to facilitate efficiency offerings to customers. This metric simply is a measure of how fast a retailer updates a customer name. It is biased toward penalising retailers over wholesalers, and has nothing to do with improving efficiency offerings to customers.	It gives an indication about how well the retailer knows its customers and how quickly it shares that info with the wholesaler through CMOS. Happy to consider other metrics for this intervention.
Provision of smart meter data to the market	Number of AMI or AMR meter readings	Is this more a target of business plans?	We wouldn't look to directly incentivise smart metering, rather incentivise the provision of smart meter data into CMOS or to retailers.

Feedback on Metrics

- UWE.7: Wholesalers ensure appropriate meter assets (including smarter meters where relevant) are sufficient to meet the needs of customers or to drive the right customer and water efficiency behaviours

Name	What is being measured	Comment	MOSL response
General comment		Role for BR-MeX somewhere here, to capture experiences of both retailers and end customers?	Add BR-MeX as a metric
Percentage of meters read	Aggregated and compared with other wholesalers	Yes, assuming this metric is referring to W reads	All reads are taken on wholesaler meters, therefore a very low rate of meter reading by all retailers that work on a wholesaler's meters is indicative of poor asset health.
Percentage of meters that are 'smart'	Proportion of AMI and proportion of AMR (including wholesaler loggers)	Business plan drivers	Indeed, that's why we wouldn't propose to have this metric mapped to a financial incentive.
Analysis on meters recording inaccurately	Number of Volumetric Adjustments where the code is Meter and the length of period this covers indicating the meter was incorrect for a long period	Somewhat already covered under MUR value that is an annual submission???	MUR is too high-level. Also this metric is looking at the period of time taken by the Wholesaler to do the Volumetric Adjustments which is useful data.
Volume of B1, B5 or C1 requests that aren't progressed or result in change of service from unmetered to assessed		No, is this an indication of wholesaler fault? Is it a good or bad thing if SPIDs are being moved from unmeasured to assessed? Not really an indication of a failure.	It is when the Bilateral process is C1 (verification of meter details - would imply that there was actually never a meter there and the wholesaler chooses to move from metered to unmeasured or assessed), B5 (Repair or replacement of a faulty meter) or B1 (indstallation of a meter)

AOB Future meetings (1 of 1)

Meeting Date	Discussion / Introductions / Endorsements
PAG 16 – 22 February 2023	1) Tools descriptions feedback on pre-reading 2) Consultation 2 outline
PAG 17 – 15 March 2023	1) TBC 2) Business Case approach & Strategic Business case
PAG 18 – 29 March	1) TBC