KPI Metrics Review of Municipalities in the Greater Toronto Area:
Existing approaches, data collection strategies, and recommendations moving forward

Corresponding Author: Dr. Calvin Lakhan

Secondary Author(s):

- **Ryan Guild**
- **Kelsey Morden**
- **Alex Christison**
- **Anita Lodi**

Prepared by: Waste Wiki@FES

Faculty of Environmental Studies, York University

Contents

1.0 Purpose: This report undertakes a comprehensive examination of key performance metrics (pertaining to the waste management sector) employed by municipalities in the Greater Toronto Area (York Region, Peel Region, City of Toronto). This report is organized into five key areas – 1) Behavioral KPIs 2) System Performance KPIs 3) Accessibility KPIs 4) Life Cycle KPIs, and 5) reuse/waste reduction KPIs. These sections include commentary on the respective benefits and draw backs of each approach, as well data requirements.

2.0 Behavioral KPIs

Defining behavioral KPIs have historically proven to be a significant challenge for municipalities, as it involves a qualitative dimension that requires gathering data directly from households. Generally speaking, behavioral KPIs in the Greater Toronto Area have tended to focus on "Customer Satisfaction" surveys – gauging attitudes towards the perceived quality of waste management services (be it pertaining curbside services, depot based programs etc.). While this approach does provide useful insight into the perceived efficacy of a city's waste management program, there are several underlying methodological issues associated with self-reporting.

2.1 Issues:

Historically, voluntary self-reporting by households tends to towards two extremes – either very positive, or very negative. Furthermore, customers who are amenable to participating in voluntary surveys already tend to have greater degrees of awareness and interest in waste management activities. It should be noted that this phenomenon is not observed when soliciting participation in household surveys through incentivization (offering renumation/gifts in exchange for participation).

A second issue with self-reporting is that participants tend to overstate their participation in recycling, or any other activity being seen as pro-social/pro-environmental. In a 2014 study examining stakeholder perception of recycling promotion and education mediums published by York University (with a focus on the Greater Toronto Area) – survey participants expressed both extremely high levels of concern and participation in recycling initiatives. However, given that these surveys were conducted in multi residential buildings (which at the time had an average recycling rate of 21%), we knew that stated actions did not match observed intent (something known as the value action gap).

Any sort of KPI work that centers on gathering self-reported data from residents needs to be interepreted with caution, as the risk of erroneous reporting can compromise the integrity of the data results.

2.2 Measurement

Quantifying qualitative feedback is both a fruitful, but challenging exercise, particularly when a municipality is not working with a prescribed methodological protocol for evaluating qualitative data.

Broadly speaking, when attempting to measure Behavioral KPI's (using measures such as attitude, awareness, normative influences etc.), two primary approaches are used: Likert Scales and Coding.

Likert Scales: A Likert Scale is a multi-point scale that offers a range of answer options — from one extreme attitude to another, like "extremely likely" to "not at all likely." Typically, they include a moderate or neutral midpoint. Most commonly in waste management research, a five point Likert scale is used to measure respondent's answers (Strongly Disagree, Somewhat Disagree, Neither Agree or Disagree, Somewhat Agree, Strongly Agree).

This approach allows municipalities to categorize survey responses, and provide “scores” to measure the intensity of participant responses.

Coding: Coding is a less common, albeit equally important technique that allows users to take open ended interview data (i.e. a conversation with a resident) and codify responses into broader categories.

While there are different techniques used for coding, the general approach is to create an “umbrella” term (i.e., separating out positive from negative statements) and then using computer software to automatically codify text based on the umbrella criteria specified by the user. Table 1 below summarizes sample coding output taken during a city on pay as throw systems in Ontario (Conducted by Dr. Calvin Lakhan, York University)

Table 1: Examples of Coding Behavioral KPIs

Do you think garbage bag limits are a good thing?	
Positive Attitudes (Yes)	Negative Attitudes (No)
"Good for the environment" – 57 “Less garbage goes to the landfill” - 34 “Reduces pollution” -15 "Promotes recycling" - 63 "Stops wasteful behavior" – 84 “Less garbage goes to the landfill” - 54	"Should be eliminated" - 97 "Inconsistent enforcement" - 55 "Unfair" - 112
Would you still recycle if your city eliminated limits on the amount of garbage bag you could put out?	
Positive Attitudes (Yes)	Negative Attitudes (No)
"I am legally obligated to" - 178 "It's the right thing to do" - 91 "It's good for the environment" - 29 "Reduces litter" - 16 "Sets a good example" - 10	"Saves me time" - 34 "Don't care" - 14 "Doesn't make a difference" - 11

2.3 Examples of Behavioral KPIs

Behavioral KPI’s (as they pertain to waste management) can largely be placed in four broad categories:

Attitudes towards the behavior: How does the participant/household view recycling (and other diversion initiatives) positively, negatively, or neutrally?

Examples: Do you think recycling is good for the environment? Do you think that recycling makes a difference?

Awareness regarding the services offered by the community:

Examples: I am familiar with which materials belong in which bin (blue bin, green bin, garbage)

I am familiar with the waste management services offered by my community. (Where is the drop off depot? Where can I purchase new carts etc.? Who do I call for bulky goods pickup?)

Normative Influence: Is there a community emphasis placed on recycling/diverting? My family and friends want me to recycle; It's my responsibility to keep my material of a landfill;

Perceived Behavioral Control: How easy it is for a household to participate in waste management program? Are there any barriers to participation (physical, financial, time) that impede participation?

Examples: Are the recycling and garbage carts easy to store during the summer?

Are recycling and garbage carts difficult to bring to the curb during the winter because of snow/ice?

2.4 What behavioral KPIs have been used to date?

Municipalities in the Greater Toronto Area (as noted above) have tended to gather data on perceived quality of waste management services. However, this has at times, peripherally included information on:

1) What encourages people to use community drop off sites/depots?
2) Frequency and duration of site visits?
3) How far are people willing to use a drop off site?
4) Satisfaction associated with Sign Signage
5) Ease of access to waste and recycling
6) Attitudes towards site layout
7) How long were you willing to wait to utilize a drop off site
8) Satisfaction associated with staff support and staff knowledge

It should be noted this is a hardly an exhaustive list. The number of potential Behavioral KPIs that can be used by municipalities depend on what component of the behavior is being examined? (I.e. attitudes, perceived convenience, satisfaction with service, awareness regarding programs etc.)

However, as mentioned in the section on measurement, while most municipalities collect behavioral KPI data in some form, they often do not engage in coding or engage in statistical techniques like factor analysis or varimax rotation to better understand what survey results are telling them.

3.0 System Performance Metrics (Including Operational KPIs)

Dating back to the early 2000s, municipal waste management performance was largely defined using three key variables: recycling rate, diversion rate and cost per tonne (gross/net). In fact, these three variables were used to develop the effectiveness and efficiency calculation that determined the Blue Box steward obligation to municipalities until the year 2010 (later amended to include Best Practice Scores).

3.1 Types of System KPIs

System Performance KPIs can largely be divided into two areas:

1) Waste Generation/Diverted/Recycled: *Note: For brevity, this applies to all waste streams collected by the municipality: Blue Box, Organics, Garbage, WEEE, MHSW, Textiles, Bulky Goods/Other recyclables, Leaf/Yard Waste
   - Total Tonnes Generated
   - Total Tonnes Recycled
   - Total Tonnes Landfilled
   - Total Tonnes Diverted
   - Kg Diverted Per Household (Both single and multi-family)
     - Including food waste reduction
   - Kg Recycled Per Household (Both single and multi-family)
   - Kg Disposed Per Household (Both single and multi-family)
     - Including avoidable food waste
   - Composition of Waste Stream (while not a KPI, it does provide useful context for how programs are evolving over time)
   - Landfill Fill Rate
   - Remaining Landfill Capacity
2) Operational Variables (including cost):
   Gross Operational Costs:
   - Collection Costs
   - Processing Costs
   - Promotion and Education Costs
   - Administration and Interest on Municipal Capital
   - Landfill Tipping Fees

   Revenue from sale of marketed material

   Net Cost = Gross Operational Costs – Revenue from sale of marketed material

- Residue Rate
- Contamination Rate (both in bin and at processing center)
- Yield Loss (At Processing Center)
- Rate of Workplace Related Injury (Drivers/Line Workers)

While these are the most commonly used metrics to communicate system performance to both internal and external stakeholders, the proliferation of light weight packaging and decreasing generation rates per capita have necessitated that alternative KPIs be explored (discussed in section X).

3.2 Data Collection

Data surrounding total tonnes generated/diverted/recycled etc. are calculated using a combination of weigh scale measurements, marketed material (for recyclables) and waste audits. Waste audits in particular pose a particular concern for municipalities, in that projections based on curbside/building audits are premised on the data being accurate. However – accuracy and reliability of waste audit data has proven to be a challenge for *all* municipalities, as there are simply not enough resource to achieve statistical significance or stratification to reasonably approximate for a broader area.

To further compound this issue, many municipalities have opted to “rotate” auditing sites from year to year (in an attempt to sample more areas of the city). While this makes sense intuitively, it actually violates rules surrounding how to correctly collect time series data, making examinations of changes over time difficult (if not impossible).

As such, while KPI measures surrounding quantities of waste generated/diverted are useful metrics when communicating the results of a program (particularly on a per capita/household basis), those results must always be interpreted with a degree of caution. Underlying deficiencies in how waste audit data is collected may compromise the integrity of any subsequent projection. It should be noted that municipalities in the GTA are now taking a more strategic approach to collecting waste audit data to improve the accuracy and predictive ability of audits moving forward. Appendix B of this report includes a brief write up providing guidance on conducting and analyzing waste audits.

4.0 Accessibility and Availability

While virtually all households in the Greater Toronto Area have access to waste management services via curbside collection (although communities also offer depot services for hazardous, electronic and bulky waste), KPIs surrounding availability and accessibility are still critical in determining overall rates of household participation.

While collection services may be offered to residents, some may not participate for a number of reasons (lack of interest, lack of knowledge regarding service schedules, lack of waste storage). This problem is particularly acute in multi residential buildings, where most households are required to bring their waste to a designated waste room or drop off location. While some buildings do provide residents with floor level waste and tri sorter recycling chutes, multi residential participation in both recycling and green bin initiatives is significantly lower than single family homes).

The most often used indicators for participation in waste and recycling programs are the set-out rate and the participation rate. The participation rate, measured from the set-out rate over a period of time, is often used to estimate the amount of recyclables and waste to be collected. However, the behavioral mechanisms that underlie the set-out rate, and hence the participation rate and the estimation of amount of recyclables collected, are largely unknown.

Although the participation rate only gives a partial explanation of diversion behavior, it provides a low-cost alternative for predicting the amount of recyclables collected and the amount of recyclables potentially collectable on an ongoing basis, if the relationship between the effective participation rate and the amount of recyclables is established. More importantly, the quantity of recyclables that is put out by less frequent recyclers, and by those who set out on every occasion, should be identified separately to investigate whether those less frequent recyclers do so because they accumulate their materials over a longer period or because they are sporadic recyclers. The sporadic recyclers participate in a recycling program on some occasions but their behavior is not consistent, as opposed to those who store their recyclables and consistently set out at variable intervals

Determining the effective participation rate largely depends on the relative storage capacity available to households. The larger the storage capacity compared with the generation of recyclables, the longer the period that should be used to estimate the effective participation rate to predict more accurately the amount of recyclables collected. While household capacity is generally not considered to be a KPI, understanding available disposal/recycling/organics capacity is critical in understanding activities such as garbage switching, illegal dumping, contamination etc. Furthermore, understanding capacity constraints within the system, particularly in rapidly growing urban communities, may indicate whether households have been provided with adequately # or size of bins/carts.

Of note, this report does not confine the term accessibility strictly in terms of "having access to waste management services", but rather, expands the term to mean "is it accessible for residents". As noted earlier, multi residential households in particular face barriers to access, even when waste management services are offered in their building.

At present, there are fewer Accessibility/Availability KPIs that specifically pertain to muti-residential buildings. Beyond infrastructural data collected by GTA municipalities (i.e. building type, waste room vs. chute, cart collection vs trailer etc.), there is no consistent way to measure whether individual units have sufficient access to waste management services. In a report funded by the Continuous Improvement Fund, convenience (or lack there-of) is often seen as the primary driver of recycling participation in multi-residential buildings. Intuitively, this makes sense – given that residents are often required to bring recyclables down to a building basement (which may be unclean, unsafe or not clearly labeled), there is an incentive to forgo participation and simply dispose of all materials in the waste stream.

Based on a review of completed CIF projects to date, increasing convenience for households (through the provision of in home recycling bags, more accessible and organized drop off points, and cleanliness of drop off points), contribute materially to observed increases in diversion noted in reports. A degree of caution needs to be taken when assuming this observed increase in diversion will persist – improved accessibility is contingent on continued efforts on both the part of households and building management to ensure that the desired behaviour continues. While households may initially be amenable to using an in home bag during the initial phase of the study, they may grow tired of having to find a separate storage space, or the time expended in doing so. Tangent to this, cleanliness and organization of drop off points requires both households and building managers to work collaboratively to ensure that accessibility is not impeded. The conditionality of this outcome highlights the necessity of prioritizing accessibility as the most critical factor for success of multi residential recycling initiatives. It requires not only ongoing participation of households with respect to source separation behaviour, but a coordinated effort to ensure that the drop off and collection of recyclables is easy to do (something that is not traditionally required of curbside single family households). Accessibility is also of equal importance to service providers, who often incur significant time costs in the event that access to recycling bins and carts is impeded in some way. Assuming that collection is provided by the municipality (or sub contracted), the additional time in collecting from multi residential households resulting from impeded access can materially contribute to elevated collection costs.

Other KPI's that measure access to waste management services include:

- % of homes with access to curbside waste collection and recycling services
- % of homes with access to green bin/leaf and yard waste services
- # of Bins per household (in non cart based programs)
- Size of bins per household
- Storage of Waste Containers (adequate space or not)

- Capacity of Waste Containers (can the provided bin/cart provide enough space for generated recyclables/organics/mixed waste)
- Density of Waste Depots (MHSW, WEEE, Mixed Waste and Recyclables) in a given geographic area, or population nexus (i.e. 1 depot per 10,000 households)

4.1 Data Considerations

While KPIs related to access and availability are largely quantitative in measure, i.e. % of homes with curbside cart collection, it is recommended that this information be supplemented with qualitative data that explore drivers of participation. "Build it and they will come", does not always apply to waste management programs, particularly for activities that require households to bring their material to a designated site or drop off.

As such, when evaluating the performance (or nonperformance) of a program, it is important to consider and compare how accessibility may be impeded (or encouraged), by behavioral attitudes and awareness regarding waste management programs. As an example, are residents aware that there is a recycling drop center in their community, or that there is a municipally sponsored textile bin program?

5.0 LCA Metrics

Life cycle analysis metrics are becoming an increasingly popular method for communicating the results and effectiveness of a municipalities waste management program. Communicating performance using life cycle metrics is gaining traction among municipalities (in lieu of weight based metrics), particularly in light of the fall overall waste tonnages (largely attributed to the shift towards light weight and composite plastics).

While Life Cycle KPIs include a wide range of potential variables, TCO2e (carbon abated), water consumption, energy consumption, emissions to air and water, municipalities traditionally focus exclusively on carbon abatement. This is largely for two reasons, 1) Existing Waste Management Life Cycle Calculators (such as the USEPA Warm Model) only communicate results expressed in terms of TCO2e, 2) The methodology for quantifying carbon impact/abatement is the most mature relative to other life cycle indicators (such as energy and water use).

5.1 Data Issue:

Quantifying the carbon impacts associated with waste management activity has proven to be an inexact science (particularly in a Canadian context). The foremost issue pertains to how one choses to define model boundaries, and whether model input data is regionalized enough to reflect the specific conditions of the market. As an example, in order to accurately model a municipalities waste management impacts, data is needed on the regional energy grid mix, total kilometers traveled for collection of waste, and transport to transfer station/MRF, knowledge of end market destinations and end use applications etc. Very few municipalities readily have access to this data. To further compound this issue, the most commonly used life cycle analysis calculator for municipal staff is based on the USEPA Warm Model, which makes assumptions surrounding transport to end markets, and includes a very limited range of materials (that don't mirror the audit categories used by Toronto, Peel or York).

At the time of report preparation, municipalities in the Greater Toronto Area have not publicly released any information regarding the carbon impacts of their waste management activities (while quantified internally, due to uncertainty regarding the results, these results have not been released).

It should be noted that York University has attempted to overcome some of the limitations associated with quantifying carbon impacts of waste management activity by purchasing a license to EcoInvent and Simapro, the world's largest commercial LCA software. Using this database, users have the ability to custom define energy grid mixes, end markets, material compositions, mode of transportation etc. to achieve a much more precise calculation.

York University has also developed a new KPI metric, $/tonne CO2e, that examines how much does a municipality have to spend on recycling activity to abate one tonne of carbon. The benefit of this approach is that it reconciles economic and environmental objectives, highlighting for municipalities what materials to target to achieve maximal carbon abatement at the lowest possible cost.

Moving forward, understanding how to best communicate life cycle analysis metrics in a way that is understandable, transparent, and methodologically sound will be critical for municipalities in both the GTA and across the country.

The waste management sector plays a unique role with respect to abating carbon, and understanding where those savings/impacts occur during the end of life management stage can help municipalities maximize environmental impact, while avoiding costs. This will become more salient if Ontario re-adopts a carbon trading scheme, wherein municipalities may be able to claim carbon credits attributable to displaced carbon resulting from recycling/diversion.

5.3 LCA Framework Assessment Halton

Halton's approach to life cycle KPIs is part of a larger evaluative framework that links service levels, asset inventories, refurbishment/replacement costs, and life cycle models to determine future capital needs requirements and gauge infrastructural risk.

Halton has developed eight life cycle models for the Public Works sector, which includes the solid waste vertical asset model (MRFs, Transfer Stations etc.)

The results of these LCA models have not been made public, however, the overall methodological design used by Halton should be seen as a better practice, both with respect to conducting an LCA, and budget forecasting for the ongoing maintenance of a waste management system.

5.4 Sample LCA indicators include:

- Collection/Transportation Fuel Use
- Collection/Transportation Emissions
- MRF/Transfer station Emissions
- Processing Emissions
- Transport to End Market Emissions
- Virgin Material Displacement Credit
- Water Consumption (Liters or Gallons)
- Waste Water (Liters or Gallons)
- Trees Saved (Total carbon abated (in kg) /22.18 (average kg's of carbon sequestered by one tree)
- Cars removed from road (Total carbon abated (in tonnes)/6 (average tonnes emitted by passenger vehicle in Ontario)

- Energy Used/Saved (KWh)
- Fuel Used/Saved (Gasoline/L)

6.0 KPIs: Measuring Waste Reduction/Reuse

To date, there has been no methodologically consistent or prescriptive way to measure waste reduction. Part of the reason for this difficulty, is that it is difficult to isolate waste reduction attributable to a programmatic change by the municipality, or whether it is part of a larger systemic trend which points to decreasing waste generation for Ontarians as a whole. Waste generation per capita is expected to decrease by 15.7% over the next 25 years. Whether this is attributable to actual behavioral change, or package light weighting, remains unclear.

As such, any proposed methodology for calculating municipal specific waste reduction has to be done in such a way that isolates between exogenous systemic change (decreasing per capita generation rates at a provincial level) and municipal specific change (causal relationships between programmatic change and a reduction in waste generation).

York University has developed a tentative mathematical approach using Log Linear analysis to isolate between the effects of exogenous (provincial) and endogenous (municipal) changes in waste generation per capita. This study is expected to be completed in December of 2018.

Of note, York University, in collaboration with Kijiji Canada, is also undertaking an initiative to measure re-use for household durable goods. Using data provided by Kijiji, York University will identify the products that are most likely to have a second life through the used good market, and identify ways that municipalities can promote re-use within their communities. This research is expected to be completed in the first quarter of 2019.

Appendix A: KPIs Used in Greater Toronto Area Municipalities

City of Toronto (Taken from Long Term Waste Management Plan KPIs, pg 91/92)

Category	Metric
Waste Strategy Related	
Waste Generation Rate	Change in waste generation rate
Single Family Residential	Kilograms/Capita
Multi-residential	Kilograms/Capita
Non-residential	Kilograms/customer
Waste Diversion Rate	Change in % diverted
Single Family Residential	Kilograms/Capita
Multi-residential	Kilograms/Capita
Non-residential	Kilograms/customer
Reduction Circular Economy – Food Waste Reduction	Change in organics in Green Bin and/or garbage bins
Single Family Residential	Kg/unit or kg/capita for units served
Multi-residential	kg/unit or kg/capita for units served
Reuse	
Textile Reuse or Recycling	Change in quantities of textiles in garbage bins

Category	Metric
Single Family Residential	Kg/unit or kg/capita for units served
Multi-residential	kg/unit or kg/capita for units served
Residual	
Residual Disposal Rate	Change in disposal rate - kg/capita/year
Other	
Customer Satisfaction Rating	Customer Satisfaction Rating
Greenhouse Gas Emissions Related to the System	Annual tonnes of CO_2equivalents reduced
Number of procurements that include waste reduction, reuse or recycling requirements such as mandating the use of recycled materials.	Total Annual
Operations Related	
Safety Performance	Annual Measures
Total Tonnage Managed	Tonnes
Green Lane Landfill Volume Filled	Cubic metres (m^3)/year
Green Lane Landfill Volume Remaining	Cubic metres (m^3)
Projected year of closure of GLL (or remaining years of capacity at GLL)	Year

Region of Peel

Waste Generation

- Total waste (all streams) generated per household
- Food and organic waste generated per household
- Performance will also be tracked on a per capita basis

Resource Recovery

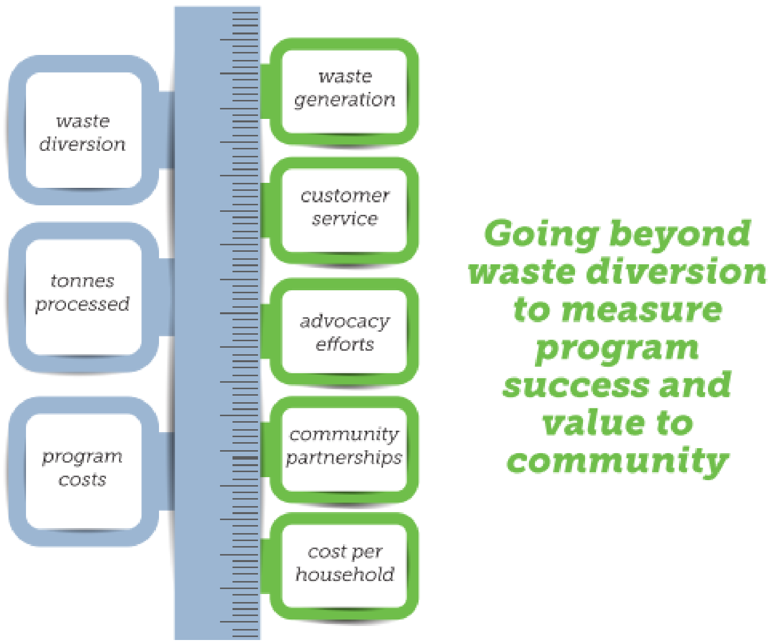
- Participation Rate by program (i.e. percentage of households participating in the Blue Box recycling program, Green Bin organics program, etc.)
- Capture Rate by program (i.e. percentage of available materials put in the Blue Box, Green Bin, etc. by residents)
- Contamination Rate by program (i.e. percentage of non-solicited materials in the Blue Box, Green Bin, etc.)
- Tonnage of reusable and recyclable goods recovered at Community Recycling Centers
- Garbage disposed of per household (Curbside and Multi-Res separately)
- Food and organic waste disposed of per household (Curbside and Multi-Res separately)

- **Customer Service**
- Percentage of Curbside and Multi-residential respondents rating collection service at satisfactory or better
- Percentage of Community Recycling Centre respondents rating Community Recycling Centre service at satisfactory or better
- Other KPIs that the Region tracks and will continue to track include financial and environmental measures.
- Financial Net operating cost per household
- Net operating cost per tonne
- Environmental Tonnes of greenhouse gases emitted as a result of residential waste operations

York Region (Taken from Smart Living Plan, pg. 40)

https://www.york.ca/wps/wcm/connect/yorkpublic/10ba5c8d-5f7f-4e49-a06b-ab270c563fb0/SM4RT_Living_Plan_Nov_28_2014.pdf?MOD=AJPERES

Figure 13 - Setting New Measures of Success



Halton Region (Taken from "Rethinking our waste" 2011-2016 plan)

It should be noted that Halton Region has not publicly updated their waste management plan past 2016. The original report "Rethinking our waste" was drafted in 2011, and as such, some of the goals, objectives and measures of performance may not be topical given the contemporary waste management landscape.

Halton's "Rethinking our waste" report did not outline specific KPI measures, but rather stated objectives that would necessitate the use of specific KPIs to gauge effectiveness.

- Enhanced Promotion and Education and Outreach (could be measured using behavioral KPIs such as household awareness, $ spent on recycling P&E per household etc.)
- Enhance Textile Communications (can be measured using behavioral KPIs to gauge household attitudes and stated participation rates in textile diversion. Measuring the quantity of textiles in the waste stream is also useful to get a sense of the size of the problem)
- Enhanced multi residential waste diversion (Measuring increases in MR recycling and decreases in contamination resulting from programmatic changes in the region.)
- Expand special waste drop off day events (Measure the increased frequency of special events, additional tonnes collected etc.)
- Expand Blue Box materials and enhance Blue Box capacity (measured as # of materials accepted in the Blue Bin, increased capacity of Blue Bins or the provision of recycling carts)
- Decrease garbage bag limits and introduce bag tags (measure reduction in number of garbage bags placed curbside in a given collection period, measure how often illegal dumping occurs, measure compliance with # of households who used bag tags for excess waste etc.)

Appendix B: Auditing consideration

While several waste auditing methodologies are employed in jurisdictions across North America and Europe, the fundamental goal(s) largely remain the same, namely: Representative sampling, and consistency of samples (sampling at the same time, same week, same location etc.). While these are certainly important considerations when conducting an audit, there exist several statistical limitations with many contemporary auditing approaches.

One of the foremost issues with auditing approaches employed by Cascadia, USEPA etc. is the erroneous use of the term statistical significance. This term is often used to describe achieving a sufficient number of samples to enable meaningful/credible analysis. However, calculating statistical significance is a mathematical exercise that denotes collecting enough samples such that unexplained variance falls below a confidence interval threshold (i.e. stated alternatively, statistical significance implies that we have collected enough samples such that the sample has a 95% probability of approximating for the actual population as a whole).

While it may not seem particularly important to make this distinction, incorrectly using the term statistical significance implies a level of precision/accuracy that is not possible through waste auditing. The number of samples required to achieve true statistical significance can number in the thousands, which is neither feasible, nor practical for a municipality/province. It is absolutely critical that any auditing methodology stress the limitations of what can and cannot be done with respect to sampling, and provide guidance regarding how to interpret/analyze the data. The goal of an audit should never be precision – rather, a sound methodological approach would be premised on "working in reverse".

The first step is identifying what resources are available and the time provided to the auditing team. The auditing team should then work collaboratively with the municipality in "placing" samples in a way that best approximates for the overall population or target group (i.e. a municipality may wish to specifically focus on large IC&I establishments).

Another potential limitation to existing auditing methodologies is that there is little prescriptive guidance regarding how best to characterize an overall population, and what data needs to be captured. Where to place audits should extend beyond simple population or tonnage based metrics – placing audits where the greatest density of people or businesses are is a short sighted approach that doesn't provide much meaningful insight into a community. What is more important is understanding the infrastructural and socio-demographic differences that exist in an area, and developing a sampling strategy that best captures the heterogeneity within a community.

Conversely, the need to stratify a sample may not even be necessary should there be minimal variance in waste compositions across areas or seasons. This consideration is also not discussed in other auditing methodologies, as an analysis of pre-existing auditing data can glean meaningful insights into how best to conduct future audits. If waste generation rates and compositions are relatively homogenous (i.e. neighborhood A, B, and C all produce similar quantities and types of waste), there is little need to sample from multiple areas. Conversely, if significant variance exists, then the need to stratify the number of samples is even more important. The ability to engage in a "pre-analysis" should data be available should be characterized as an auditing best practice, as the findings can save significant time and money for municipalities looking to develop an auditing strategy moving forward.

Proposed Sampling Strategy for Single Family Waste Audits

Step 1: Calculating the number of samples required

When developing any sampling strategy (be it waste audits, conducting surveys etc.), the number of samples required will be a function of satisfying statistical validity requirements, and budgetary/resource constraints.

With waste audits in particular, the costs are sufficiently prohibitive that a municipality will never be able to collect the number of samples required to gather a representative sample size (which is normally measured at the 95% confidence interval, +/- 5%).

With this in mind, a proposed approach is to "work in reverse", where in we estimate the total number of audits the budget allows for, and then allocate those samples to specific housing types/geographic regions. The goal is to "place" these samples in areas that serve as a rough approximation for the municipality as a whole.

Step 2: Allocating the Samples

Once we have determined the number of samples that the budget allows for, the next step is to allocate the samples to account for the following factors:

- Different types of households
- Geographic proximity
- Demography
- "One Offs/Aberrations"

Step 2a) Accounting for different types of households

The breakdown of single family households in Municipality X is as follows:

1. Single-Detached Houses: 78,975 (67.81%)
2. Semi-Detached Houses: 20,240 (17.38%)
3. Row House: 17,215 (14.78%)

4. Other Single Attached Homes: 35 (0.03%)

As such, our samples should be allocated to approximate for the relative contribution of each housing type. Assuming that our budget allows for 20 audits, we would want to sample 14 Single-Detached Households, 3 Semi-Detached Households and 3 Row-Houses. Given that the "Other" housing type constitutes such a small percentage of overall households, there is a rational for omitting it from the data set.

Step 2b) Accounting for Geographic Proximity and Demography

Once we have calculated the number of samples we will take from each housing type, the next step is to sample households from different geographic regions within the city.

For the purposes of simplicity, let's assume that Municipality X can largely be divided into four areas (North, South, East and West).

While there is an initial inclination to simply allocate the number of samples to each area evenly (5 from each area), consideration needs to be given to population density. As an example, if we know that East Municipality X is home to 60% of the population, then we would need to weight it accordingly when allocating samples.

- As an illustrative example, let's assume that the population breakdown of Municipality X is as follows:
- East Municipality X: 60%
- North Municipality X: 20%
- South Municipality X: 10%
- West Municipality X: 10%

Therefore, of our 20 samples that the budget allows for, 12 should be conducted in North Municipality X, 4 in North Municipality X, 2 in South Municipality X and 2 in West Municipality X.

The next step is to overlay this data with what we have calculated from Step 2a. The number of samples for each housing type, in each region (accounting for population density) is:

	East	North	South	W
Single Detached	8.14	2.71	1.36	1.
Semi Detached	2.09	0.70	0.35	0.

Row House	1.77	0.59	0.30	0.

Note that we do not have rounded figures for samples – While the above describes a largely quantitative exercise, selecting areas for audits is also a qualitative judgement as well. These are rough guidelines for how to allocate samples – the project manager has a certain amount of latitude in terms of where they want to conduct audits.

Step 2c) Accounting for One Offs

Extending upon my earlier point surrounding qualitative judgement, there are certain areas of the city that we may want to sample because we know that there is something peculiar about it, i.e. Springdale Single Family homes may actually have 3 families living inside, or areas with a high incidence of illegal basement apartments.

I would recommend that 2 samples be used to target these “One off” areas, to get a better understanding of how these a-typical housing arrangements affect household generation and recovery.

Step 3: Comparing Samples from Previous Audits

Whenever you are comparing audits taken at different times (to assess trends in waste composition etc.) we must follow the “like with like” principle.

The following criteria should be used when comparing audits taken at different times:

- Must be the same housing type
- Must be the same geographic region
- Must be the same season
- Only the relative composition of the waste/recycling streams can be compared

The last point may require some elaboration, in that conventional auditing procedures would compare the weights of materials disposed/recovered. However, if the Bin type (or frequency of collection) has changed, then you can no longer conduct comparisons using weight based metrics. Only the relative contribution of the waste stream can be readily compared.

It is also recommended that an analysis of variance test be conducted on all samples to determine the degree of variability among samples taken from different regions, housing types etc.

This is an important test, in that if we observe a low variation in waste/recyclables composition between study sites, then there is no need to stratify the audit samples moving forward. We operate under the assumption that different housing types/areas produce different quantities and types of waste – this assumption may not be valid. Households in Municipality X may actually have homogenous consumption and disposal habits.

Appendix C – Analytical Concerns

Data Source & Sampling Issues

How was the data collected and from where? Was there a sampling methodology used when selecting audit areas?

Why this is a concern:

In my experience, samples taken during waste audits rarely (if ever) approximate for the municipality with any sort of statistical significance. This is largely due to the resource and time constraints of having to take the necessary number of samples to satisfy statistical significance requirements at the 90 or 95% confidence interval. The subsequent implications of having a data set that does not adequately approximate for the total population ranges from relatively innocuous (we can still make some good guesses) to severe (you cannot move forward with any sort of analysis, as the underlying data doesn't really tell us anything).

Data Consistency Issues

This can actually be considered as an extension to source and sampling issues, but were samples taken from the same households/streets over the duration of the audit data?

Why this is a concern:

Conventional wisdom would probably be no, as most people think that by sampling different areas each time, we get a better "snapshot" for the city as a whole. However, from a strictly statistical standpoint, this is actually the exact opposite of what you want to do.

One of the great things about collecting data over time, is that we can then perform "Time Series Analysis" (allowing us to calculate trends over time, and control for exogenous changes to the overall system using fixed or random effects regression). Unfortunately, the moment you no longer sample from the same areas, our data goes from being a "time series", to being discrete snapshots at yearly (or however often audits are conducted) intervals. This severely restricts the type of analysis you are able to do, as (from a strictly statistical standpoint) you cannot measure changes over time.

Unfortunately, there is little, if anything you can do to overcome this issue – at best, you can make "good guesses" about trends and changes over time, but they cannot be substantiated using any conventional statistical procedure.

Events in Isolation or Multiple Overlapping Events?

Are program changes (switch from bins to carts, weekly pick up to bi-weekly) carried out in isolation (one change is made in a year) and city wide (whatever is changed, is done across the entire city at the same time)?

Why this is an issue?

It's important to highlight the difficulty of isolating the effects of programmatic changes. There are two reason for this:

- You need to do something called log linear analysis to determine the strength of the relationship(s) between programmatic changes and any changes in diversion/waste composition etc. This isn't a clear cut procedure, as it is predicated on a number of data assumptions that may not be satisfied (and going back to the previous point, if you don't have time series data, you can't do log linear analysis).
- There is an inter-temporal dimension to programmatic changes – any change to the system may not be realized until several years into the future (if at all). You have to define appropriate time scale boundaries to gauge the effect of changes.