

Trip report – Niger final survey preparations

Niamey, October 27 – November 10, 2019

This report trip will be a more concise version compared to my other trip reports because of my current workload.

October 27 to November 3

At the beginning of this week, I touched base with the team on the progress of the preparation for the 24-hour recall. We revised the household questionnaire to make it lighter and more efficient and revised surveys tools on Tuesday and Wednesday before the team left to do their first recalls on Wednesday night. We got the team back on Thursday, Friday and Saturday to debrief, retrain, update tools and correct mistakes. The team went out again on Sunday.

We made the critical decision at this point to no longer use CSDietary for data entry (please see note below). We are going to use a simple data entry program linked to the FRIL file with the numeric codes for foods and recipes. This is the same extensive FRIL that Jikatt and Aichatou had worked on and that was also used by the surveyors for the standard recipe data collection.

November 4 to November 10

I went out with the team on the field to supervise how the second phase of the field test was going. We came back to Niger on Thursday to do another round debriefing, retraining, and tool updating.

I also met with Yatta to put in place a plan to continue working on the databases in preparation for data processing and analysis. Balarabé and I briefly discussed the possibility of meeting again for continued technical support on data processing and analysis, but we first needed to see data collection end and take stock of where we are at that point.

Note sur la saisie des rappels des 24 heures

Mourad Moursi

5 Novembre 2019

J'avais promis à Balarabé de regarder de près aujourd'hui la situation des différentes tables de données (table de composition des aliments, table des recettes et table des facteurs de conversion) et d'évaluer leur état d'avancement pour la saisie avec le logiciel CSDietary. Le logiciel démarre complètement vide d'information et on importe dedans les différentes tables spécifiques au pays, d'où l'importance de leur état d'avancement.

Pour rappel, ce logiciel permet de faire immédiatement le lien entre la saisie des fiches papiers et les différentes tables pour convertir les quantités en grammes et faire le calcul des ingérés en temps réel au fur et à mesure de la saisie. Dans les enquêtes de rappel des 24 heures, ce processus de traitement des

données prend typiquement 3 à 6 mois avant de pouvoir avoir les premiers résultats sur les ingérés totaux. Le logiciel permettait de réduire ce temps à presque zéro en transférant les efforts nécessaires à avant la collecte de données au lieu de faire ce travail après.

Après examen détaillé des tables, et quelques essais de saisie, je ne pense pas que nous puissions utiliser le logiciel CSDietary. Il nous faut changer de stratégie.

Les tables ne sont pas à un état suffisamment avancé pour pouvoir les utiliser dans le logiciel. Bien que le logiciel présente de nombreux avantages, il a un énorme inconvénient : tout ce que ne se trouve pas déjà dans ses tables importées ne peut pas être saisi. Donc si le superviseur en charge de la saisie se retrouve par exemple avec une recette non listée dans les tables ou bien s'il manque un facteur de conversion ou une méthode de mesure pour un aliment ou une recette, il est obligé de s'arrêter. Il devra faire une sauvegarde partielle de la fiche en attendant que l'administrateur des tables puisse faire la mise à jour et l'envoyer à tout le monde sur dropbox pour téléchargement. Etant donné les difficultés rencontrées pendant le pilote avec la saisie et la mise à jour (en l'absence de connexion 3G fiable) d'un simple questionnaire sur les tablettes, il est presque impossible d'envisager quelque chose encore plus compliquée comme la gestion en temps réel de 3 tables différentes avec remontée d'informations du terrain de 15 superviseurs différents pour savoir ce qui manque. Il faudra aussi que les superviseurs assurent une gestion exemplaire des fiches partiellement saisies et savoir à tout moment quelles sont celles qui sont manquantes. C'est un scénario qui va frustrer tout le monde, pour les superviseurs qui seront constamment interrompus dans leur saisie comme pour le gestionnaire des tables qui sera sous une énorme pression surtout les 2-3 premières semaines.

Par conséquent, il nous faut changer de stratégie et je vous propose une approche qui avancera sur deux fronts simultanément :

- 1- **La saisie principale** : il nous faut développer un nouveau masque de saisie des fiches de rappels et des fiches recettes qui ne sera rattaché à aucune table et offrira donc aux superviseurs plus de flexibilité dans la saisie. On pourra extraire certaines informations rapidement de cette saisie comme par exemple le nombre moyen de lignes remplies par personne par rappel pour participer au contrôle de la qualité ;
- 2- **La saisie randomisée de contrôle** : en parallèle, je vous propose d'installer le logiciel CSDietary sur une ou quelques machines au niveau central. Cet effort consistera à échantillonner des fiches au hasard par région et par équipe pour les saisir sur le logiciel CSDietary au niveau central.

La saisie principale

Cette nouvelle approche de saisir plus « librement » et sans contrainte de tables existantes entrainera automatiquement quelques changements résumés ci-dessous :

- Il faudra ajouter des colonnes dans les fiches rappel et recette pour pouvoir mettre des codes numériques aux :
 - Aliments et recettes
 - Méthode de mesure
- Il faudra développer une liste de codes pour les aliments et recettes. Fort heureusement, grâce au travail déjà effectué par l'équipe avant et pendant la standardisation des recettes, nous avons déjà une telle liste. Mais il faudra passer du temps, **cette semaine avant le début de la collecte**

des données, pour harmoniser les codes avec ceux que nous avons déjà utiliser pour les recettes standards et avoir une liste finale de codes qui sera utilisées par les superviseurs sur le terrain ;

- Il nous faudra aussi coder les méthodes de mesure, mais comme celles-ci sont fixes, ce sera très facile ;
- Les superviseurs vont coder manuellement tous les aliments et recettes sur les fiches avant de faire la saisie.

La saisie randomisée de contrôle

Comme indiquée ci-dessus, cette saisie sera effectuée pour un nombre limité de fiches tirées au hasard parmi les fiches de l'enquête. Il faudra néanmoins mettre en place un système d'échantillonnage qui permettra de tirer des fiches de chaque région, équipe et enquêteur pour couvrir tout le monde de façon systématique tout au long de l'enquête.

Pour ce système puisse vous être utile, il vous faut envisager de saisir au moins 1 rappel par enquêteur par semaine. Cela veut dire 60 fiches à saisir par semaine. Une personne peut raisonnablement saisir environ 15 fiches par jour sur CSDietary, soit donc 4 jours de travail. Mais évidemment si 2 personnes se relayent par exemple, cela ramène le nombre de jours à 2 jour par semaine dédiés à la saisie.

Je vous conseille très fortement de mettre cela en place tout au long de l'enquête en portant une attention toute particulière aux deux premières semaines. Evidemment, plus vous pouvez faire de fiches et relayer des gens sur la saisie, plus vous aurez un meilleur contrôle sur la qualité.

Ce système présente aussi un autre avantage considérable : au fur et à mesure de la saisie de contrôle, il va falloir mettre à jour les différentes tables. Même si vous n'allez plus utiliser ces tables pour la saisie principale, la saisie de contrôle vous permettra de les avoir prêtes (j'estime à plus de 90%) pour l'analyse statistique avec Stata (ou autre logiciel) de vos données. Cela réduira très fortement le temps et l'effort à dépenser après la fin de collecte pour traiter les données. L'équipe a déjà eu un avant-goût de la difficulté à traiter les données avec les recettes standards qui, à ce jour, 3 mois après la fin de la collecte des données, ne sont toujours pas finalisée malgré le nettoyage et les ateliers de traitement des données. Les données des rappels des 24 heures avec des dizaines de milliers de lignes à traiter seront encore plus complexes et difficiles. A part les données elles-mêmes qu'il va falloir nettoyer et vérifier, ces données vont devoir être synchronisées avec 3 autres tables, et cet effort-là n'est pas du tout à sous- estimer. Il vaut mieux donc répartir cet effort sur le temps en se forçant à travailler sur les tables pendant les 38 jours de collecte de données à travers la saisie de contrôle.