

January 9, 1984

LOS ANGELES: SCHOOL OF PUBLIC HEALTH

To: Charlotte G. Neumann, M.D., M.P.H.
 From: Anne H. Coulson
 RE: KENYA TRIP REPORT (12/6/83 - 12/23/83)

Discussion with Mr. Njeru - December 12 1983:

Problem learning TPL - Mr. Chana reluctant to teach.
 Joe Mambo at ICL know, hasn't been able to schedule course for Njeru, others.
 All four scope faces go with same format - one floppy disk.
 Can use 80 columns for now; will need to rerevise forms for final formatted input.
 Njeru has no experience with variable length files - only rectangulars.
 Maurice Njundo is an operations source.
 To transmit from floppy to computer, we will need special time (computer cannot accept MAXIMOP at same time as transmit) before real start up Mondays. This seems ok.
 Need operator. This is alright with Njeru. Otherwise we must compete with Mrs. Kinanjui at lunch break.
 Njeru has passed training - needs keys.
 We need to arrange for systems engineer in case of Saturday or Sunday breakdown - loss of time, esp. paid time, plus backlog of work.
 Could arrange? policy
 retainer
 fee for service
 combination
 Njeru is trained to take care of your average hang up, but not into the electronics, etc.
 Our office is in Geography Bldg., Rm. 215, phone: 334244 ext. 2321. It is to be used for files of originals that come for key entry and are retained here.
 Shelving will be installed.

Meeting with Professor Scott - Coulson, Carter, Martin, Njeru, Kagia:

He is going to London 17-12-83 for 15 days.
 The computer is being upgraded -
 Virtual Ap System & JPMS - for 32 bit architecture
 Can use full ASCII 8 bit 9 track system
 MAC will replace MAXIMOP
 Version 9 of SPSS can be installed also GENUSTAT, MINITAB, BLIMIII
 All this should be in place by end of March.

Compatibility problems - Mr. Pius Mwangi will look into this immediately, I bought unlabelled EBCDIC from U.S., Mr. Mwangi will try to translate and use. Mr. Mwangi will also create a tape in ASCII on ICL, translate to EBCDIC and ask CBS to try to read on 370. If it is OK, I will receive a copy to take to LA to try on our machine. Mr. Mwangi met with us and was advised by Professor Scott that this is of first priority. Dr. Scott and Mr. Mwangi claim they "do this all the time." (Personally doubtful about latter - hopeful it will work.)

Assuming compatibility works--

Keys-Dr. Scott will arrange for immediate keys to Njeru - there are four - front door, mid-door, data entry room and outer data entry room. We were advised of costs we would incur if keys lost (locksmith, four locks, new keys for other users. We can start this weekend.

Use of facilities-Dr. Scott would prefer that we be out of day use of facilities after January 1st - restrict to weekend use. We have assigned locked files there for our staff. Room should be locked up and tidy for Monday morning.

Transmission of data-floppy to mainframe. Dr. Scott agreed that this could be done at 7:30 each Monday morning. Operations start at 7:00 am - if the machinery is ready, could do at 7:15 am. Dr. Scott suggested Njeru come at 7:15 am or 7:30 am.

Service-Dr. Scott called ICL about service policy for that machinery (key entry, disk) for weekends. They will call him back. (He is to be in touch with me before leaving.)

Dr. Scott also suggested that we could use the IDS terminals (key to disk with their own disk) in case of breakdown at Chiromo. (Eric prefers not to get involved with another institute). Dr. Scott suggested Njeru and I see Dr. Kinanjui about this possibility. I plan to do so. (Eric may be right, but I am here now. To make provisional arrangements, or at least to meet Dr. Kinanjui cannot be all bad - in case of need the groundwork has been laid).

Personal ID devices - This is needed to access the key to disk equipment. Mr. Akuta will provide this for Mr. Njeru before this Saturday.

TPL

Professor Scott expressed surprise and annoyance that Mr. Chana has not arranged to teach Mr. Njeru ICL and in fact seems reluctant to do so. He said that we would see to it that Mr. Chana did as he was told and that he would be told to spend his first two days back after his leave teaching Mr. Njeru. (Might be a good idea to have Bill learn as well - see if this can be arranged.) Mr. Chana will be back right after 1/1/84. This should be in time to be of value.

After this discussion, Mr. Njeru was excused at Dr. Kagia's request. Dr. Kagia wished to discuss possible training for Mr. Njeru. Dr. Scott said that a Post Graduate Diploma program had been approved through the University but that the "part time" aspects were still a problem. There is a possibility that Mr. Njeru could audit the PG Department course currently offered - scientific and engineering computing - and another in the spring - statistical programming and computing. Each of these would be four hours per week in class. They would provide good training for the project but there are two problems - first, it may not be possible to spare Njeru for the scientific class hours and the prep. time in the busy upcoming period. Second, Dr. Scott fears that an advance "peek" at the courses might give Mr. Njeru an unfair advantage in the examinations. (I shall write him

Charlotte G. Neumann

January 9, 1984

Page 3

I will communicate further with Scott and will send memo to understanding to him before I leave.

Discussion regarding Data Management Activities

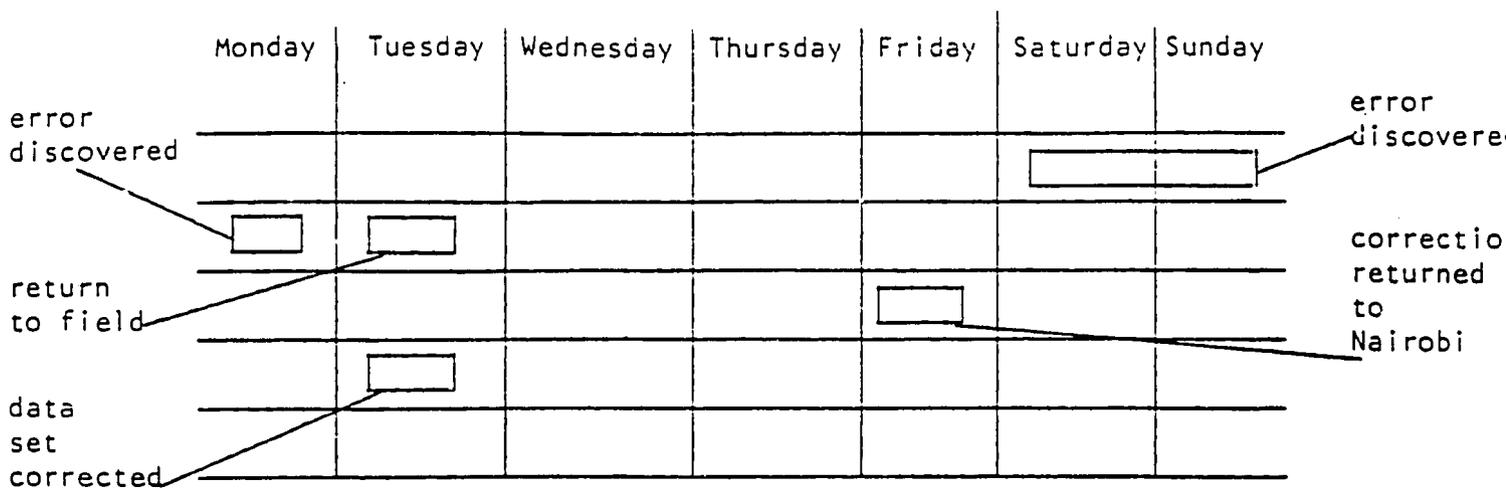
Following our discussion with Dr. Scott, Dr. Kagia brought up a number of very useful points concerning the data management plans. He was particularly concerned with the flow of data from Embu to Nairobi and the back flow of connections with subsequent return to Nairobi of connected materials. His points were well taken and resulted in three steps being taken:

1. Mr. Njeru will not be the data messenger but will be primarily based in Nairobi. Mr. Martin will meet with him at least once per week in Nairobi to assess progress and handle problems. Mr. Njeru will continue to report - via mail - to Mrs. Coulson on a bi weekly basis, through Mr. Martin.
2. Mr. Njeru will be requested to write a program that will edit on the computer the data input and will print out error statements (We, Martin, Coulson, and Njeru, later met on this matter).
3. For the first six weeks of the study, errors will be circled in red by key entry operators and returned to the field in that manner. These will be errors of out-of-range (absolute) and of inability to read, or for two operators to agree on the reading. There may be some "recognition of problem" by the key entry operators (e.g. nonsmoker with 2 pack/day recorded) but I would not expect this benefit so early in the study.

This procedure will require TPL but will not require computer range checks. It has a positive side as well, in that it relates the field enumerator and the key entry operator in their tasks and returns to the enumerator his or her individual work that is somehow in error. The learning curve for the enumerators, using this procedure, should be quite steep, so that when the error messages replace the instruments with red error marks, there should be few of them, and they should be better understood.

One further item was discussed - errors and correction time. It was felt that correction time would be item specific. If it is "too late" to correct an item, it will be specified as missing if there are other measures of the same item (height, for example). If the error is in the "only collection" of the item, the material must be collected again, and perhaps the whole of that instrument repeated. This will be worked out as soon as the instruments are completed.

Errors are to be corrected or declared "missing" within two weeks. That is, given an error discovery on a weekend and report back to the field on Tuesday .



In exceptional cases, more time may be required. If so, a special error message should go to Nairobi, and the record concerned should be transferred to hold-file, awaiting correction.

USAID justification for stop in London

I. Contact with Mr. Shah at Mayfield Ltd.

1. Authorized given number of disks (30) and tapes (5). Gloria will confirm by telex. Stopped paper shipment, can buy from Compurite in Nairobi.
2. Mr. Shah will send advance shipping copies to Mr. Taleeb at Compurite who will handle customs, etc. He will also send copies to Eric Carter. Purchaser is UCLA for delivery in Nairobi.
3. Mr. Shah says it is not ready to go, given authorization and confirmation on disks, ribbons. He assures me that it should arrive before Christmas.

II. Contact with ICL, makers of the computer at University of Nairobi (and makers of the peripherals as well). There are three basic problems and one training question

The underlying problem is the need for IBM (or similar US) compatibility with the ICL produced tapes. There are three ways this could be accomplished:

- a. the current ICL 2950/20 with a DME2 DS could be forced to do the job by some simple means inherent in the system or achieved easily, based on the experience of others.
- b. the current ICL and/or its DS and/or its programming or peripherals could be modified to allow compatibility. This might be a new OS, special software, etc.
- c. the current ICL could be used to produce ICL compatible tapes which

4

could be translated to IBM readable at ICL in London, the new tapes would be sent to UCLA, the old back to Kenya. Tapes for the Kenyans would follow the reverse route. If all of these options fail, we better rethink the IBM compatible service bureau option. The training question is one for Njeru. His stature would be immeasurably increased by going "somewhere" for further training. ICL might have some short courses which we could enroll him in - one very useful one would be formatting the input for 7502 or a fuller understanding of SPODE, or variable length file management.

These were the problems I took to ICL. I first called from the US and was transferred from secretary to secretary. I was unable to achieve a direct appointment. On my arrival, I called the various components of ICL.

I. Can ICL stuff be IBM readable.

see problem notes -

if yes (absolute! - no questions) and conditions of compatibility are feasible (upgrade cost, translation cost, time cost) go on to item IIa. There is uncertainty, go on to IIb. If the answer is no either to compatibility or feasibility, go on to IIC.

IIa. Proceed with University of Nairobi plans -

Key to tape, transfer of
computerized data, etc (see IIIa below)

IIb. Proceed to test points of uncertainty, trying out system. I brought an unlabelled tape from our Amdahl (not IBM but fully compatible) which would demonstrate if IBM - ICL compatibility exists. ICL-IBM could be tested with a tape I take back, or by transfer of tapes between the Gort computer (IBM) and the University ICL. Any other points of uncertainty should be appropriately tested. If OK go to IIa above. If not, go to IIC below.

IIC. Make arrangements to transfer key entry operations from the University to a service bureau with IBM compatible key entry. (Comp-rite?) This is a most unfortunate alternative - for Kenyan interface it will require utilization of the govt computer or other IBM compatible installation and will seriously retard capability for training, interaction, etc. Key entry operators will be unnecessary; Mr. Njeru's position will either be deleted or radically downgraded to a "smart" messenger service to the service bureau (I plan to check into the availability and costing of such bureaus this week). If this is the alternative, see IIIb below.

5

- IIIa. Assuming key entry at University of Nairobi, it is essential that Mr. Njeru learn to prepare the formats for the key entry machines. Without this capacity, the key entry operation is subject to considerable error, is substantially slower, and loses the possibility for missed item check and single variable range check inherent in the programmed format. When I was here in August, Professor Scott and I had an amicable meeting in which we agreed that Mr. Njeru would receive this training in September. It is now December, he has not received the training and it now appears that he must take Mr. Chandra's course. If he "does well," he will then be permitted to program the formats. This can be the occasion of substantial -and intolerable - delay for the project, to say nothing of increased expense and error potential. If this can be resolved so that:
- a. Njeru can learn from Mr. Chandra now, with sufficient personalized instruction to make him competent, or
 - b. Mr. Chandra can be persuaded (clout, money, status, whatever) to do the formatting for us - it is largely a one time job for the completed forms, needing little reworking but some continuing oversight (perhaps Njeru could do that,) or
 - c. Njeru could be just through an ICL course on the subject (either here or in London) so that he would achieve competence and perhaps have a bit more standing at the computer center.
- IIIb. Key entry by service bureau will involve the service bureau doing everything from key entry through range checks and rejection of unacceptable values. It may be possible for them to do computer editing as well, but this may be costly. They would then have the key entered material transferred to magnetic tape; one copy would be retained here, one sent to UCLA where further processing is done. Tapes with corrected, linked or merged data will be sent back from UCLA to Embu for whatever use can be made of them.

There is an alternative which could start immediately after the completion of key entry and corrections. That is to connect to the Government computer center - which is IBM. If time can be obtained there, then all of the functions planned for the ICL at Chirromo can be transferred to the government's IBM.

Specs on Problem

1. To the extent possible we wish to use and support the computing facility at the University of Nairobi. This includes key to disk data entry, creation of the computer tape file, computer range checks, linkage and merging of files with consistency checks, error correction creation of files for analysis and analyses. While it would be desirable to be able to use the University computer for all of these functions, it is essential that it be used to create a computer tape file from the key to disk entry of data.

2. The computer file must be IBM compatible - that is, at a minimum, the tape file created when the disks are entered into the computer must be placed on a magnetic tape in such a format, etc., that the tape can then be read on an IBM 3033 or other similar machine.
3. It is also desirable that tape files created on the IBM be readable on the University computer so that data available only through the US computer can be added to the files at the University computer or so that, should the linkage, merge and error correction steps prove unfeasible here, completed data files for analysis could be returned to U. Nairobi.

There are four possible ways this necessary compatibility could be attained:

1. The ICL 2950/20 with its present operating system may be able to create and use IBM readable tapes.
2. The ICL 2950/20 may be upgraded (at reasonable cost) to achieve the two way compatibility. This may involve hardware, software, US or programming modifications.
3. The ICL 2950/20 may be able to produce tapes which can then be converted by some other ICL machine - at, naturally, a service fee - into an IBM readable format. Presumably the reverse conversion would work to produce ICL readable tapes. This would have to be handled using ICL technical assistance.
4. The same arrangements of intermediate conversion may work using other computers as intermediaries. This could easily be done using the DEC PDP 11, VAX, Burroughs, or other American computers. Again, there would be an assumption of service cost, and that reversing the procedure would produce suitable ICL tapes.

If the compatibility cannot be achieved through any of these - or other - means then we must reorganize the Kenyan base of data management transferring to CBS or a service bureau that is able to do this. The data management for the project cannot go forward without assured compatibility.

It would be desirable for the compatibility to be assured, based first on compatibility experience within ICL or among ICL users. This assurance would then be tested by exchange of tapes ICL to IBM and IBM to ICL. Failing some experimental background, assurances based on theoretical grounds could be fully tested. Such tests would need to be more exhaustive.

Clearly, the simplest, least disruptive and least costly alternatives should be used.

Up until now, the data management package for the project has been predicated on the utilization of the U. of Nairobi computer at Chiromo. This installation is an ICL 2950 with a DME operating system, terminal accessed through a system called MAXIMOP. Batch access is also possible. Key entry facilities include four key to disk terminals accessing one eight inch floppy disk. Key entry can be formatted using a special language TPL (Terminal Programming Language) This language is not especially difficult but does have some "tricks." Currently, Mr. Chana is the only person at the facility conversant with the language.

As a result of our discussions with Dr. Scott, Director of the facility, in August we had arranged that we would use the key to disk machines on weekends and possibly some evenings as well, as it was felt that our work would swamp the facility's "service bureau". We would need to hire our own key entry staff which was done, and Mr. Njeru would have to learn to use the key entry machines in order to act as key-entry supervisor. The actual instruction for training Mr. Njeru was not issued until late October (two months after the agreement) and his training, except for TPL, was not completed until late November. Mr. Chana has totally resisted any of our efforts with regard to his training Mr. Njeru in TPL.

It is essential that the data tapes created in the field be readable in the IBM configuration available at UCLA or in the IBM compatible configuration at UCB. We have been working on this problem, trying to arrange for the ICL computing facility to read tapes created on the IBM and to write us a supposedly compatible tape to try at UCLA. The Data content of either tape is unimportant - it is the readability that is to be determined. Actually, two tapes have been taken to U. Nairobi, both contain two files of preliminary survey data. The first, which I took in August was an IBM Standard Label; the second, which I took in December, was unlabelled. Efforts to read the first tape were unsuccessful; efforts to read the second tape were ordered on a priority basis by Dr. Scott. The staff member so directed did not, apparently, have the expertise, required unavailable documentation (for the process, not the data set) to do the job and had not started the job one week after he was directed to do so (the week was to allow the running of the tape while I was still in Nairobi). He was also instructed, at an even higher level of priority, to create an IBM readable tape from any available ICL processed data, have it tried on the IBM 370 equipment at CBS and, if it worked, to give me a copy to bring home. The tape was not created as late as Wednesday evening, December 21st, again one week after the directive was issued. We therefore have no direct evidence of compatibility except Dr. Scott's report that lots of other people have done it.

To try to discover something about the basic compatibility problem, I stopped in London to visit ICL headquarters. It was my belief that they should be able to talk sensibly about the general aspects of IBM-ICL compatibility, whether or not it was possible when the U. Nairobi ICL configuration, whether additional hardware/software would make it possible, or whether other ICL computers, could accept the 2950 material and convert it to IBM compatible form. This would of course, entail a service fee, if it could be done at all. I found ICL particularly unhelpful in this regard in London, and at their referral, in Nairobi. No one that I talked to in the company could give me any assurance of compatibility.

In talking to others about the U. Nairobi facility and use of "foreign" taped data sets, I found a belief that the process was difficult, with indifferent success. These informants included persons who had lost data or had been unable to use data brought from the US. No one I talked to had had total success with the compatibility, though all had been interested only in the IBM to ICL compatibility which is, at present, of lesser interest to us.

Given these considerations, it seems unreasonable to rely on Dr. Scott's assertions regarding compatibility. There is no indication that we will be any more successful in getting test tapes to assess this for ourselves than we have been over the past year. I do not feel that, under these circumstances it is feasible to rely on ICL input of study data.

The set up of our own "service bureau," using the university equipment nights and weekends is not working out well. Admittedly, this has been a training period, but the use of persons hired full time for key entry on only Saturday and Sunday, with fill in work in Nairobi to cover the rest of their 44 hour week is extremely expensive for the work produced. We have had only one weekend session, as Mr. Njeru was not considered qualified to have the keep until December 16th. Because Mr. Chana has been unwilling to teach Mr. Njeru the TPL, only an 80 column entry program could be used. The group - Mr. Njeru and four operators, worked on Saturday, December 17th (Mr. Martin was able to stop by to see them). They continued to work on Sunday, but engineers working on the computer managed to activate the building circuit breaker, cutting out all power for a considerable time. When power was restored, the floppy disk content of one and a half days work of five people proved irretrievable. These circumstances make it appear not feasible to try to run our own service bureau operation, especially with the nights and weekends restriction and the requirement that the key entry operators be full time employees.

There are three other supporting arguments for the decision. The first is that Mr. Njeru is not really a competent key entry supervisor, nor is this the best possible use of his time. The second is that we will have to pay for a preventive maintenance - emergency repair contract for the floppy disk system if we are to be assured that it will be usable for the full weekends. The cost of this is unknown at present. The third is that the university will probably charge us a connect fee for the use of the machinery. This has not come up yet, but well may, considering the charging history for this project. (Our formal request to be considered a university research project and to receive the benefit of the research rate was denied.)

With the compatibility and key entry problems, it seemed advisable to seek other sources of key entry services. Accordingly I talked to the data people at AID and with a number of commercial service bureaus in Nairobi. These included Comprete, Data Center, Computer Consultants Ltd and Software House. On the basis of discussion with others, the last named was deleted from consideration. Results of discussions with the other service bureaus are summarized below:

Data Center - This organization is large, well thought of. It uses ICL exclusively and no one there was willing to speak to the compatibility use. The person one should speak to there about data services was on leave, and no one else could discuss any questions. This seems an unlikely contender.

Comp-rite - These are the Apple dealers. They have a service bureau operation and will provide us with a price. They would be willing to enter into a contract but do not feel that they could offer what we need. They recommended that we talk to computer consultants, ltd.

Computer Consultants - I met with the managing director and his assistant. They are interested in our business and would like it all. They are IBM trained and use IBM machinery, supporting a service bureau staff of key entry operators, programmers, etc. They enter onto floppy disks and would produce IBM standard label magnetic tapes to send to the U.S.

Their maximum costs, which they indicated a willingness to negotiate, based on volume and duration of work are well within our key entry budget and, in fact, will cost less (according to Bill Martin) than we are presently paying for our four key entry operators. Their charges would include key entry machinery, operators and supervision; they will correct their errors, free of charge. The Manager expects, based on Eric's number estimates, that a person/week will be required for key entry and verification. (The two procedures will be done by different individuals, of course.)

Data Flow -

Data comes from field (laboratory) on computer compatible forms. Data is from observations, from responses, from measurements and from laboratory determinations. Each individual form will include the information collected by one enumerator, team of enumerators, lab tech or laboratory. For example, all questionnaire responses and physician observations and measurements (e.g. blood pressure) will be recorded on one or two instruments. The fact of specimen collection will be included, but the laboratory will be sent a special form to record its measurement. This means that completed data will not be held up (or be at risk of loss) waiting for completion of tests or for other information.

Data collection instruments will be completed by enumerators and turned in to supervisors. The supervisors will be responsible for checking over forms for completion and as necessary, feasible and/or useful, for special problems the group or the enumerator may be having with from completion. The supervisor will note any problems and will call for appropriate action by enumerator; e.g. clearer numbers, completion of all items, etc.

After checking back with enumerators on problems, forms will be forwarded to the Embu office. Any further coding or other processing (depending on form) will be done there. The forms will be logged in as completed on the study management file (Apple). If the form is carboned, the original will be detached for transport to CCL and key entry. The carbon form will be retained for safety while the original is in process. If the form is not carboned, the form will be copied before being sent for data entry; the original will be sent and the copy retained for safety.

Completed forms - originals - will be logged for shipment to CCL. Key entry operators will red circle unacceptable values and unreadable values as they key enter. If there are no corrections on a form, it will be processed to forward to U.S. If there are corrections to be made, forms will be corrected in the office, if possible; returned to supervisor who will correct form, query the enumerator or arrange recollection of the information, if feasible. If the item cannot be corrected by any of these means it will be recorded as "missing" for that collection. Corrected forms, with corrections in green (or some other constant ink color different from blue, red, or black) will be returned to the Embu office and sent back to CCL.

Returned forms will be checked off on the log as "returned, ship to U.S." or "returned for correction." Corrections will be relogged out when sent back to CCL and will be checked off as above for returned forms.

Original forms will be sent to U.S. weekly or biweekly (depending on volume). Before shipment, any corrections will be recorded on the carbon (or xerox) form so the two sets are the same. (Extensive corrections may be more easily recorded by copying the form or the corrected part of it instead of copying by hand).

For the first three months of the study, all forms will be sent to the U.S. by airmail. If the rekeying demonstrates that CCL key entry is adequate then 10% of forms will be sent by airmail from month four on. (There may be adjustments, up or down, in the 10% figure, depending on rekeying results). UCLA will send a sealed list of forms to be sent airmail which is to be opened only when forms are ready to send, to determine if particular forms are to be sent air or surface mail. Airmail will be sent on the weekly or biweekly schedule established. Surface mail may be sent on a monthly basis.

Forms on which UCLA has questions will be identified - if possible, the form number and question number will be sent with the query. If necessary, a copy of the form will be sent.

Forms from the field, laboratory, or other facilities within Kenya will flow through Embu and be key entered through CCL. Forms with data originating outside Kenya will be processed through UCLA. In the latter case, carbon or copy will be returned to Embu, along with tape copies. The content of the tape copies - linked, etc. - will be determined. (At present a full tape copy with all data, linked, etc, is planned for every three months).

Tape copies of data (corrected) will be sent on a regular basis to the U.S. - at least once each month. Two tapes will be made, one will be sent. The floppy disks containing the data will not be recycled until UCLA has verified receipt and reading of the tape. Then the floppies can be reused.

This whole seems rigid - I realize that changes will be both desirable and necessary. Let's put this in place as the general description of flow for now - and as a basis for discussion and mutually agreed upon change.

Njeru duties -

With the new data management arrangements, Mr. Njeru's duties will be substantially changed. He should be responsible, under Bill Martin's supervision, for transport of the forms to CCL, for interaction with CCL about corrections, etc (more message than policy, please) and for funnelling of corrections at Embu and in the field. As he has no Nairobi responsibilities other than those described here, he will have time for these duties. He may be made responsible, for the logging in and out of forms or other activities related to data management. I suggest that he not be assigned to a supervisory role involving calculations or whatever - I would prefer that he be wholly involved - at whatever level of work is needed - with the "completed forms and on" data management under Bills' supervision. He should continue to give Bill biweekly (or semi-monthly) reports would appreciate receiving a copy.

As to his computer use - Professor Kagiã is anxious for him to start analysis, if only of the preliminary survey. No harm there. There are two things on the computer that Njeru could do effectively between now and summer:

1. Work on the problem of making the IBM format work on the (new) ICL (no point in working on the old) use copies of the tapes sent to US, or tapes I brought in August or December)
2. Get acquainted with the CBS computer (an IBM 370 or something)

I suggest for the ICL computer, that Njeru be given a small - specified - amount of computer funds and a very small working time allowance - renewable as needed on detailed application to Bill Martin who will decide on his knowledge and in consultation with Eric and Anne.