medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

So Much to Do: Expected Duration of Initial Consultations for Sick Children in Low- and Middle-income Countries Using the Integrated Management of Childhood Illness Clinical Algorithm

Anthony Carter[1]

1 Emeritus Professor of Anthropology, University of Rochester, New York, United States

Abstract

Background: Few studies have examined the time required for primary care consultations; none have focused on sick child visits in LMICs. This paper begins to fill that gap by providing evidence-based estimates of the time needed for initial visits with under-five infants and children at public or not-for-profit facilities in low- and middle-income countries that have adopted the Integrated Management of Childhood Illness (IMCI) clinical algorithm.

Method: Estimates of the mean expected duration of IMCI consultations are based on (a) classification profiles, i.e., tabulations of the health issues presented by patients less than 5 years old; (b) lists of the tasks included in applicable versions of the IMCI algorithm and the conditions that elicit them, and (c) an estimate of the time needed to perform tasks with no pre-defined minimum duration. The latter requires, in addition to classification profiles, information on rates of task performance and the mean observed duration of consultations. I found such information by searching MEDLINE, the database of the International Network for Rational Use of Medicines, the websites of the WHO and its regional offices, GOOGLE, and GOOGLE SCHOLAR using search terms such as 'Integrated Management of Childhood Illness',

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

'observational', 'prospective', 'classification', 'clinical signs', 'health facility survey', and 'validity'. I also reviewed studies that cited a qualified study and, conversely, material included in the bibliographies of qualified studies.

Results: I have found or constructed classification profiles for 38 patient populations in 22 countries. The mean expected duration of consultations ranges from 16.3 to 18.7 minutes in the infant populations and from 15.9 to 26.3 minutes in the child populations. Largely because of mandatory tasks, the distributions of expected consultation durations are left-truncated. The shortest consultations range from 14.9 to 16.7 minutes in the infant populations and from 10.4 to 16.8 minutes in the child populations. One in 50 patients requires a consultation that, depending on the mix of cases in the patient population, exceeds the mean duration by 13.2 to 32.1 percent.

Conclusions: IMCI consultations with infants and children presenting for the first time with a new concern require considerably more time than previously recognized. This confirms that the quality of most IMCI consultations is low and that the personnel costs of providing appropriate IMCI services are substantial. The steps required to produce estimates of the expected duration of IMCI consultations suggest that surveys designed to measure the effects of IMCI training must include tabulations of the gold standard classifications found in the patient populations.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

So Much to Do: Expected Duration of Initial Consultations for Sick Children in Low- and Middle-income Countries Using the Integrated Management of Childhood Illness Clinical Algorithm

Introduction

This paper provides a series of evidence-based, case-mix adjusted estimates of the time required to provide initial consultations for sick children in public primary care facilities in low- and middle-income countries (LMICs) that use the Integrated Management of Childhood Illness (IMCI) clinical algorithm.

Though health care providers have a great deal to do during primary care visits – begin the consultation, establish rapport, elicit the patient's concerns, gather information about presenting illnesses, perform a physical exam, discuss diagnosis and treatment, summarize, and bring the consultation to a close [1, 2] – little is known about how much time all this requires. In general, it is thought that consultations rarely should be shorter than ten minutes. In the United Kingdom, where the average consultation lasts about nine minutes, the Royal College of General Practitioners considers that 10-minute consultations are adequate only for "patients with simple needs" [3]. The Primary Health Care Performance Initiative (PHCPI), an organization concerned with the quality of care in LMICs, regards 10-minute visits as "the lower limit of what is sufficient" [4]. It also is thought that longer consultations generally are superior to shorter ones. The PHCPI recognizes that effective communication, "appropriate diagnostic questions, adequate physical exam," and an "appropriate treatment plan" are not guaranteed by longer

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

visits [4]. Nevertheless, during longer consultations, clinical officers and assistant medical officers in Tanzania ask more questions and perform more examinations [5]. Similarly, during longer consultations in Europe and North America, physicians recognize more issues, explore more psychosocial concerns, give fewer prescriptions, and/or provide more preventive measures [6-13]. Longer consultations often are essential for those with multiple morbidities [3]. Finally, several experienced physicians in the United States and Argentina regard 15 to 20 minutes as the optimal duration of primary care visits [14, 15]. The time required for IMCI consultations is thought to fall at the low end of this range. A leading member of the team that designed the IMCI algorithm, suggests that IMCI consultations "can be completed in 10-15 min although some cases take longer" [16]. According to the World Health Organization (WHO), the mean duration of sick-child consultations in primary care facilities where the IMCI algorithm is used is 15 minutes [17]. However, none of this work takes account of the mix of problems with which patients present [12, 18], the tasks involved in care for patients with these problems, or the time required to perform such tasks. And no evidence is presented to support the claims regarding the duration of IMCI consultations.

The IMCI clinical algorithm and the research surrounding it provide unusually rich resources with which to address these issues. A comprehensive program to reduce "unacceptably high" infant and child mortality in LMICs, IMCI was developed in the mid 1990s by the WHO and the United Nations Children's Fund (UNICEF) [19]. IMCI seeks to encourage proper health practices in families and communities, to strengthen health care systems, and to improve the technical quality of primary care services. The IMCI clinical algorithm is the key means used to achieve the last of these aims and has been adopted by more than 90 countries

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

[20]. Concerned with the co-occurrence of issues involved in the bulk of child morbidity mortality in resource-poor settings [16, 21, 22], the algorithm integrates a variety of interventions shown to be cost-effective in previous vertical programs [23]. For infants less than 2 months old, the algorithm focuses on bacterial infections, feeding problems, low weight, and, in some versions, jaundice. For children 2-59 months old, the foci include acute respiratory infections, especially pneumonia; diarrhea; fevers, especially malaria and measles; malnutrition, and anemia. Immunization status is a concern for both age groups. The algorithm provides a scheme to classify the health issues with which infants and children present, an array of tasks providers may be expected perform, and criteria by which tasks are elicited. Research on the design and utility of the algorithm, its effects on provider performance, and related topics furnishes data on the frequencies of gold standard IMCI classifications in a variety of patient populations. In some cases, it also enables one to calculate the time required to perform tasks.

METHODS

The expected duration of IMCI consultations in any given setting is a function of the health issues presented by patients less than 5 years old, the kind and number of tasks health care providers are expected to perform, and the time required to perform them. The health issues of patient populations are indicated by classification profiles. Similar to, but broader than, clinic-based morbidity profiles, classification profiles are tabulations of the IMCI classifications and other indicators of health status found in a patient population. The tasks required are those described in the applicable version of the IMCI clinical algorithm. Estimates of the time required to perform tasks that have no predetermined duration are based on studies of IMCI services for

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

which there is information on the health issues presented, rates of task performance, and the mean observed duration of consultations.

Definitions

Classifications. IMCI classifications are categories used to organize signs and symptoms indicative of the occurrence and severity of conditions that account for a substantial portion of child mortality in low- and middle-income countries. Classifications differ from "etiologically distinct diagnoses" in two ways [24]. Some combine distinct diseases into a single category. "For example, a child with the classification 'very severe febrile disease' could have meningitis (viral or bacterial), … sepsis (from a number of possible pathogens), … cerebral malaria," pertussis, plague, or pneumonia. Conversely, a single disease may be assigned different IMCI classifications if it presents with different levels of severity. If pneumococcal pneumonia presents with cough and fast breathing it is classified as 'pneumonia'; however, if it presents with cough and stridor it is classified as 'severe pneumonia/very severe disease'. Similarly, in areas of high malaria risk, malaria presenting with fever alone is classified as 'malaria'. When it presents with fever and a stiff neck or unconsciousness it is classified as 'very severe febrile disease.'

IMCI classifications are grouped into sets of mutually exclusive alternatives. All versions of the algorithm include a set of classifications for 'diarrhea' divided into subsets for 'dehydration', 'persistent diarrhea', and 'dysentery'. The initial generic algorithm for infants <2 months old also provided sets of classifications for 'possible bacterial infection' and 'feeding problem or low weight'. Subsequent India algorithms for infants added a set of classifications for 'jaundice'. In addition to the classifications for 'diarrhea', the initial generic algorithm for

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

children 2-59 months old provided classifications for 'cough or difficult breathing', 'fever', 'measles', 'ear problem', 'malnutrition', and 'anemia'. Subsequent national versions have added 'throat problem', 'dengue' as an aspect of 'fever', 'HIV status', and 'mother's HIV status'.

Tasks. A task is a piece of physical work IMCI-trained providers may be expected to perform during a consultation. The performance of a task requires providers to use their ears, eyes, hands, and/or voices in ways that consume time. The IMCI clinical algorithm is an arrangement of tasks.

Providers also are expected to make a variety of decisions, as in 'Classify COUGH OR DIFFICULT BREATHING', 'Decide the Malaria Risk: high or low' if a child has a fever, and 'IDENTIFY TREATMENT'. However, in this paper, these decisions do not count as tasks because they are cognitive rather than physical and are not inherently time-consuming. Providers may sometimes take time over decisions because the issues are novel or complex or because they are uncertain or confused, but, in principle, such decisions can be made instantaneously.

IMCI tasks vary in terms of the conditions under which they appear, their complexity, and their minimum duration. Eliciting conditions divide tasks task into three groups:

- Mandatory tasks must be performed for all patients in the relevant age group. These consist of greetings and other preliminaries with which consultations begin, a series of assessments, a check for immunization status, and a check for problems that fall outside the scope of IMCI.
- Conditional assessments are triggered by the appearance of certain signs or symptoms during mandatory assessments or in the initial series of conditional assessments.
- Treatments and counseling are triggered by prior classifications and other findings.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The simplest IMCI tasks involve just one step. In the case of a child with diarrhea, for example, 'look for sunken eyes' as a sign of dehydration. More complex tasks involve multiple steps. For example, 'give an appropriate oral antibiotic' involves, in addition to a decision concerning the correct drug(s) and dosage, a minimum of 6 and as many as 7 steps:

- 'Tell the mother the reason for giving the drug to the infant.'
- 'Demonstrate how to measure a dose.'
- 'Watch the mother practice measuring a dose by herself.'
- 'Ask the mother to give the first dose to her infant.'
- 'Explain carefully how to give the drug, then label and package the drug.'
- 'Explain that all the oral tablets or syrups must be used to finish the course of treatment, even if the infant gets better.'

And, if more than 1 drug is required,

- 'collect, count and package each drug separately.' [25]

Several mandatory tasks and conditional assessments have predetermined minimum durations. To assess an infant's breastfeeding, a provider must "ask the mother to put her infant to the breast" and "observe the breastfeed for" a minimum of "4 minutes" [25]. Determining if a patient has fast breathing requires counting her breaths for at least 1 minute and repeating the count for another minute if the initial count is high [26]. Taking a patient's temperature with a thermometer requires at least 2 minutes [27]. The time required to perform other tasks is not specified and must be estimated separately.

Data

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The data used in this paper come from two sources: (a) observational studies of clinical encounters involving IMCI-trained health care providers and children less than 5 years old presenting for the first time with a new concern and (b) handbooks and chart booklets containing versions of the applicable IMCI algorithm.

An observational study is one in which a researcher examines each patient and in which, if information on the performance of health care workers is provided, a researcher is present during, and takes notes on, each clinical encounter. I have located such studies by searching MEDLINE, the database of the International Network for Rational Use of Medicines, the websites of the WHO and its regional offices, and GOOGLE SCHOLAR using as search terms 'Integrated Management of Childhood Illness', 'IMCI', 'observational', 'prospective', 'classification', 'clinical signs', 'health facility survey', and 'validity'. I have expanded these searches by locating and reviewing studies listed by GOOGLE SCHOLAR and other resources as citing a qualified study and, conversely, material included in the bibliographies of qualified studies.

IMCI handbooks are designed "to help teaching institutions incorporate IMCI into academic programmes for doctors, nurses and other health professionals" [26]. Chart booklets are intended to be used as mnemonics by providers while they are engaged in caring for patients. I have located applicable IMCI documents using GOOGLE and GOOGLE SCHOLAR. Most of the world is covered by one or another generic version. Only a relatively small number of countries have produced their own adaptations. The key links between observational studies and relevant IMCI documents are the country in which, and the years during which, the observational data were obtained.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Analysis

Classification Profiles. I construct classification profiles for observed patient populations in several ways. Where a study reports standard IMCI classifications, these are used directly. Where researchers report disease diagnoses but use their local knowledge to provide equivalences between disease categories and IMCI classifications, I follow their lead. Where researchers report disease diagnoses with clear IMCI equivalents together with the incidence of cases requiring hospitalization, I treat those requiring hospitalization as severe and those not requiring hospitalization as moderate. For example, among infants less than 2 months old, hyperbilirubinemia requiring hospitalization is IMCI 'severe jaundice', while hyperbilirubinemia not requiring hospitalization is IMCI 'jaundice'. Where researchers report the incidence of clinical signs as part of an effort to test their value in IMCI algorithms for infants <2 months old, I convert signs to classifications by following the applicable IMCI algorithm. (Classification profiles for observed populations appear in columns A-D of the worksheets in Workbooks S2-S5. For additional details on the construction of classification profiles, see Text S1.)

IMCI algorithms provide negative classifications only for 'feeding problem or low weight' among infants and 'ear problem', 'malnutrition', 'anemia' and, where relevant, 'symptomatic HIV infection' among children 2-59 months old. I derive other negative classifications from observed data. For 'cough or difficult breathing', for example, the available classifications are 'severe pneumonia or very severe disease', 'pneumonia', and 'no pneumonia: cough or cold'. To these I add the negative classification '≠ respiratory disease'. In the WHO's Eastern Mediterranean Region, the classification 'no ear infection' is given to children whose caretakers respond affirmatively to the question, 'Does the child have an ear problem', but in whom the provider

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

can find none of the relevant clinical signs. Providers are to advise the mothers of children with this classification 'to go to ENT specialist for assessment'. To avoid confusion around this issue, I use the classification '≠ ear problem' for children whose caretakers respond negatively to the initial screening question. Because negative findings such as '≠ respiratory disease', '≠ fever', etc. are equivalent to 'none of the positive classifications', the incidence of negative findings is equal to 1 minus the sum of the observed probabilities of the positive classifications in the relevant set.

All versions of the IMCI algorithm for children 2-59 months old include treatment and counseling tasks to be performed only for patients less than 2 years old. The KwaZulu-Natal version also includes treatments whose use depends on whether the patient is less than 1 year old. Complete classification profiles for children 2-59 months old, therefore, should include the incidence of children who are less than 2 years old and, in sites where something like the KwaZulu-Natal algorithm is used, the incidence of children less than 1 year old. Since children <6 months old are not at risk of dengue fever, profiles for populations where this is a concern should include the incidence of this age group.

A complete classification profile for both infants and older children also includes the incidence of patients who lack 1 or more immunizations and the incidence of children with health problems that fall outside the purview of IMCI.

None of the profiles include classifications for which there are no reports of incidence. Among these are 'wheeze' for children in Morocco and South Africa. I also exclude the incidence of 'general danger signs' for children 2-59 months old. The incidence of these signs is rarely

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

reported and, since they play a role in subsequent assessments and classifications, they are redundant.

Combinations of Classifications and Other Determinants of Care. Sick infants and children very frequently suffer from 2 or more health issues. Moreover, health issues themselves and the forms of care to be provided for them may interact. As a result, the types and numbers of tasks required for the average patient in a population is best estimated by means of a list of all possible combinations of IMCI classifications and other findings that trigger conditional assessment and/or treatment and counseling. In constructing such lists, I assume that classifications and findings from each set of mutually exclusive items co-occur randomly. In the worksheets in Workbooks S2-S5, the lists begin with column AA. Each set of mutually exclusive IMCI classifications or findings is allocated a pair of columns. The left-hand column in each pair contains the specific finding; the right-hand column contains probability of the finding. The lists exclude sets of classifications not used in the relevant studies or for which no patients presented with the relevant clinical signs. It should be noted that positive classifications for persistent diarrhea and for dysentery do not occur unless diarrhea is present and that positive classifications for measles and dengue do not occur unless fever is present.

Column Q in the worksheets of Workbooks S2-S5 shows the probability of each possible set of classifications. This is the product of the prevalence of each item in the set. The sum of the values in column Q is shown in the last nonblank cell of column N. This figure should be 1.00. However, because EXCEL has limited numerical precision, the actual figure usually is slightly less, especially where the number of combinations is very large.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

IMCI Tasks and the Conditions of Performance. The clinical tasks included in this analysis are drawn from the original IMCI handbook or subsequent chart booklets. Clinical tasks and the conditions under which they must be performed are indicated in these documents by a subset of sentences using imperative verbs that direct providers to carry out specific activities. Nearly all these sentences occur in Chapter 2 and Parts II-III of the IMCI handbook and in the sections of chart booklets concerned with assessing, classifying, and treating infants <2 months old and children 2-59 months old. (For a detailed description of the standards used in identifying tasks, see Text S2.)

Numbers of Tasks Required. I obtain the mean number of tasks required in each patient population by applying the performance conditions built into the applicable version of the IMCI algorithm to the combinations of classifications and other determinants of care described in the relevant classification profile. Tasks are excluded from the analysis if the need for their performance in a population cannot be determined.

Columns S-Z in the worksheets in Workbooks S2-S5 worksheets show the number of tasks required for each combination of classifications and other findings: mandatory tasks to the left, conditional assessments in the middle, and treatment and counseling tasks on the right. Tasks with predefined minimum durations are shown separately from those without.

The figures for mandatory tasks come directly from relevant imperative sentences in the applicable versions of the IMCI algorithm.

The number of conditional assessments with no minimum duration required for each combination of classifications and other indicators is given by EXCEL formulas with the following form:

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

= *23-IF(AB4="≠ Cough/Difficult Breathing", 4)-IF(AD4="≠ Diarrhea", 6)-IF(AH4="≠ Fever", 6)-IF(AJ4="≠ Measles", 3)-IF(AL4="≠ Ear Infection", 4)+IF(OR(AN4="Very Low Weight", AP4="Anemia"),3)*

In these formulas, expressions such as 'AB4' refer to cells in worksheets of Workbooks S2-S5 showing the mutually exclusive sets of classifications and other indicators of health status found in a patient population. The version shown here is based on the 1995-2005 generic algorithm. Twenty-three conditional assessments with no minimum duration are required for a patient who has 'cough or difficult breathing', 'diarrhea', 'fever', 'measles', and 'ear problem'. That number is reduced by 4 if the patient has no respiratory disease, by 6 if she has no diarrhea, and so on. To this result, 3 conditional assessments are added if the patient is classified 'very low weight' or 'anemia'. Similarly, the number of times providers must perform a 4-minute breastfeeding assessment for an infant less than 2 months old in a population using the 2003 India algorithm

= *IF(AND(AB4<>"Possible Serious Bacterial Infection",AD4<>"Severe Jaundice",AF4<>"Severe Dehydration",AH4<>"≠ Feeding Problem"), 1, 0).*

That is, a breastfeeding assessment must be performed if and only if the infant has a 'feeding problem' but has not yet been found to require referral, i.e. has not been classified 'possible serious bacterial infection', 'severe jaundice', or 'severe dehydration'.

Each positive classification or finding, plus the negative classifications 'no feeding problem' for infants <2 months old and 'no ear infection' for children 2-59 months old, calls for the performance of one or more treatment or counseling tasks. If a patient has more than one classification triggering the same task, the task should be performed only once. If a patient is

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

given a classification that requires urgent referral to hospital, critical pre-referral treatments followed by referral pre-empt all other treatment and counseling tasks. In several of the algorithms for children 2-59 months old, the rule governing the occurrence of 'give a course of an appropriate oral antibiotic' therefore is

= IF(AND(OR(AB4="Pneumonia", AH4="Dysentery", AN4="Acute Ear Infection"),BC4= 0), 1, 0).

Here the expression 'BC4=0' indicates that urgent referral is not required. The formula indicates that a patient is to be treated with an oral antibiotic if and only if she does not require urgent referral and she has one of the following classifications: 'pneumonia', 'dysentery', or 'acute ear infection'.

In a few instances the effect of multiple classifications may be greater than the sum of its parts. For example, according to the India version of the algorithm, infants who have received the classifications 'some dehydration' and 'severe persistent diarrhea' require urgent referral even though neither classification triggers urgent referral on its own.

Counts of treatment and counseling tasks appear in two locations. The columns on the right side of the worksheets in Workbooks S2-S5 indicate whether or not a specific treatment or counseling task is required by each combination of classifications and other determinants of care. Tasks triggered by findings that are not routinely reported – the occurrence of other problems, missing immunizations, and, for children 2-59 months old, age – appear at the extreme right of the worksheets. Columns X-Z in the worksheets for infants (Workbook S2) and W-Z in the worksheets for children 2-59 months old (Workbooks S3-S5) show the total number of treatment and counseling tasks expected for each combination of classifications and other findings, the first for core classifications and the others for 'other problems', missing

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

immunizations, and, for children, counseling regarding the feeding of a well-nourished child <2 years old. These figures are the sums of the figures found in the relevant columns on the right of the worksheets in which occurrences of each task are reported separately.

Minutes per Task with No Specified Duration. The information needed to estimate the time required to perform the average task with no minimum duration, M_{nd},, is available only for 7 population of patients 2-59 months old. In these populations, M_{nd}

$$= \Big(\big(D_{obs} - ((T_{cb} * P_{cb}) + (T_{tt} * P_{tt} * 2))\big) \div \big((T_{m_{nd}} * P_{m_{nd}} * K_m) + (T_{c_nd} * P_{c_nd} * K_c)\big) + U_{ab}\Big) * Q$$

where the variables are defined as follows.

D_{obs} is the weighted mean observed duration of consultations in minutes.

T_{cb}, T_{tt}, T_{m_nd}, and T_{c_nd} are the weighted mean number of times providers are to count breaths for 1 minute, take a temperature with a thermometer, and perform mandatory and conditional tasks with no minimum duration. Failure to perform mandatory tasks is implicated in failure to perform conditional task [28], so the 2 groups are treated separately.

P_{cb}, P_{tt}, P_{m_nd}, P_{c_nd} are the performance rates for counting the breaths in 1 minute, taking a temperature with a thermometer, and performing mandatory and conditional tasks with no minimum duration. If P_{cb} or P_{tt} are unavailable, P_{c_nd} takes the place of the former and P_{m_nd} takes the place of the latter.

U_{ab} is the proportion of patients who receive an unnecessary antibiotic.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

K_m *and* K_c, are the proportions by which the performance rate for mandatory and conditional tasks with no minimum duration are reduced for studies in which the results are inflated by conspicuous observation, and

Q is the proportion by which the time required to complete a task with no minimum duration must be reduced for studies in which the results are inflated by conspicuous observation.

Measuring task performance is complicated by several factors. To begin with, task performance may be, and often is, affected by the presence of an observer. It has been suggested that observer effects are short-lived [29, 30]. However, to minimize the risks of exaggerating the time required to perform tasks with no minimum duration, I assume that observer effects may be ignored only when providers were not aware that observers were present or when researchers show that the presence of an observers had no measurable effects. In all other cases, the performance of tasks with no minimum duration is corrected for observer effects using 3 coefficients. K_m adjusts performance rates for mandatory tasks, K_c adjusts performance rates for conditional tasks, and Q adjusts for the pace of task performance. The values of these coefficients are derived from Rowe et al.'s [28] comparison of simulated client and conspicuous observation surveys in Benin (see Workbook S7 and Text S3). Where no adjustment for the observer effect is required, the value of each coefficient is 1.

Tasks also may be, and frequently are, performed incorrectly [27, 31]. On the one hand, required tasks may be performed poorly and/or with critical steps omitted. A child may be weighed while fully dressed. A child's respiratory rate may be counted inaccurately, when she is agitated, or for less than a full minute. An oral antibiotic or antimalarial may be prescribed but the dose, frequency, and number of days in which it is to be administered may be incorrect or

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

omitted. A mother whose child does not require urgent referral to hospital may be told some but not all the signs indicating that her child should be brought back to the facility immediately. On the other hand, providers may perform tasks unnecessarily. Because of the longstanding concern with rational use of antibiotics, the most frequently measured error of commission is unnecessary use of oral antibiotics. To account as fully as possible for the time required to perform tasks with no minimum duration, I take unnecessary as well as required tasks to have been performed if their central elements or a minimal set of such elements is reported to have occurred. For example, the task 'give an appropriate oral antibiotic' is counted as having been performed whenever an oral antibiotic was provided even if it was not indicated and even if the dose, frequency, and duration are incorrect or omitted.

Portions of some tasks are performed outside the confines of a consultation. For example, Plan B rehydration for patients with 'some dehydration' involves "an initial treatment period of 4 hours" during which "the mother or caretakers slowly gives a recommended amount of ORS solution … by spoonfuls or sips" [26]. Rehydration is to take place in a separate place in the facility, presumably while the provider works with other patients. Plan B rehydration is preceded by a determination of how much ORS the patient requires and a series of instructions to the caretaker concerning the way the ORS should be given. It is followed by a reassessment of the patient and further action as required, either additional Plan B rehydration or Plan A. None of the studies reporting the mean duration of consultation explain how they handle Plan B. However, it is clear that, in the few instances in which it might be required, they measure only the preceding instructions and concluding assessments.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Mean Required Duration of Consultation. For any given population, the expected duration of the average IMCI consultation, D_{exp},

$$= \left(4 * T_{bf}\right) + \left(2 * T_{tt}\right) + T_{cb} + \left(M_{nd} * \left(T_{m_nd} + T_{c_nd}\right)\right)$$

where

T_{bf} is the number of times a provider working with an infant must perform an assessment of breastfeeding;

T_{tt}, T_{cb}, T_{m_nd}, T_{c_nd}, and M_{nd} are as defined in the previous section.

RESULTS

Study selection

I identified 47 observational studies of primary care consultations for children less than 5 years old presenting for the first time with a new concern that report gold standard IMCI classifications, clinical signs, and/or disease diagnoses. Nine of these studies had to be excluded. Two use disease categories without indicating how these map on to IMCI classifications [32, 33]. One reports the prevalence of several IMCI symptoms but not the relevant classifications [34]. Three fail to report mild classifications for two or more symptoms and/or to distinguish among severe, moderate, and mild classifications [35-37]. Two studies report on very small subsets of classifications [36, 38]. One reports classifications for persistent diarrhea and dysentery but fails to mention dehydration [39]. A study from Nigeria [40] appears to assume that no child has more than 1 gold standard classification.

Study characteristics

Relatively complete classification profiles can be constructed for 38 studies carried out between 1993-94, when the IMCI program was just getting underway, and 2014-17 (Workbook

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

S1). Nine were in India; 4 in Bangladesh; 3 in Kenya; 2 each in Afghanistan, South Africa, and Tanzania, and 1 each in Benin, Bolivia, Botswana, Brazil, Burkina Faso, Cambodia, China, Egypt, Ethiopia, Gambia, Ghana, Morocco, Sudan, Uganda, Vietnam, and Zimbabwe. Since the Cambodia study included infants <2 months old as well as children 2-59 months old, the 38 studies contain data on 39 distinct patient populations.

I identify each patient population with an expression consisting of the standard three-letter abbreviation of the name of the country in which it was carried out, a 2-digit number to distinguish among studies done in the same country, and either 'I' or 'C', depending on the age group observed. For example, the 1993-94 IMCI validation survey in Kenya is identified as 'KEN_01_C'.

Twelve of the patient populations consist of infants <2 months old. Twenty-seven consist of children 2-59 months old. For infants, the number of consultations observed ranges from 34 to 1268. For children the range is from 98 to 2724.

The original generic IMCI algorithm, in use between 1995 and 2005, applies to 7 of the 12 infant populations. The 2003/2009 India adaptation applies to the remaining 5. Excluding 'other problems', the 1995-2005 generic version covers all the health issues seen in 14 of the child populations and most of the issues seen in the Botswana, Cambodia, and Vietnam patients. The 2002 KwaZulu-Natal provincial adaptation applies to ZAF_02_C. National adaptations apply to AFG_01_C, AFG_02_C, EGY_01_C, IND_01_C, and MAR_01_C. The 2008 generic version applies to BFA_01_C and CHN_01_C. I also assume that the 2002 KwaZulu-Natal adaptation applies to HIV status in BWA_01_C and to all classifications in ZWE_01_C and that the Timor-Leste adaptation applies to dengue in KHM_01_C and VNM_01_C.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

In 30 of the classification profiles, all the classifications and findings come directly from the studies used. IND_01_C and IND_03_C were constructed using local knowledge of the equivalences between disease diagnoses and IMCI classifications. BOL_01_I, GHN_01_I, IND_02_I, IND_03_I, KEN_01_I, and ZAF_01_I were constructed, in part, by applying the relevant IMCI algorithms to reported clinical signs. BOL_01_I, GHN_01_I, IND_02_I, IND_03_I, and ZAF_01_I also required using information on the need for referral. Figures for 'severe malnutrition' and 'very low weight' in TZN_01_C come from a contemporaneous household survey in the same region. (For additional details on the construction of classification profiles, see Text S1.)

Eight of the profiles are substantially complete: IND_06_I and IND_08_I for infants and BRA_01_C, EGY_01_C, IND_01_C, MAR_01_C, SDN_01_C, and UGA_01_C for children 2-59 months old. Figures for 'persistent diarrhea' are missing for BEN_01_C, KEN_02_I, and TZN_01_C. Figures for 'dysentery' are not available for BEN_01_C or KEN_02_I. Figures for 'measles' are missing for BEN_01_C, TZN_02_C, and ZAF_02_C. Figures for ear infections are not available for BFA_01_C. In addition, the source for ZAF_02_C does not report the incidence of anemia and the source for BWA_01_C, where HIV/AIDS is a concern, does not report the incidence of 'mother HIV-positive'. In TZN_01_C, 22 children required urgent referral, but the specific severe classifications are not provided.

No information on the incidence of 'other problems' is available for 7 of the child populations. The figures for 'other problems' in CHN_01_C (3 conditions), ETH_01_C (7), GMB_01_C (8), IND_01_C (16), and IND_03_C (10) appear to be based on complete and fully specified enumerations. The enumerations for the following profiles are not fully specified:

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

AFG_02_C ('skin rashes/allergies' and other), BEN_01_C (1 and other), BGD_01_I (6 and other), BGD_02_C (1 and other), BOL_01_I (11 and other), BRA_01_C (asthma and other), GHN_01_I (10 and other), IND_02_I (14 and other), IND_04_I (8 and other), IND_05_I (7 and other), IND_06_I (8 and other), IND_07_I (6 and other), IND_08_I (7 and other), ZAF_01_I (9 and other). The figures in the following profiles refer to specific conditions, but likely are incomplete: EGY_01_C (1), KEN_01_C (1), KEN_02_I (1), MAR_01_C (2), SDN_01_C (1), VNM_01_C (2). The studies on which the figures in KHM_01_I, BGD_03_C, BGD_04_C, and UGA_01_C are based provide no specificity.

Figures for immunization status are provided for 2 of the 12 infant populations and 11 of the 27 child populations.

The incidence of children <2 years old is available for 17 of the 27 child populations.

Of the three child populations in which HIV is a major concern, mother's HIV status, one of the triggers for assessment of a sick child for symptomatic HIV, is available only for ZAF_01_C. The incidence of children <1 year old, a condition for the use of co-trimoxazole to prevent HIV transmission is available only for ZWE_01_C and ZAF_01_C.

Information on the observed duration of consultations and task performance rates comes from the studies used to construct the profiles for BRA_01_C, KEN_03_C, MAR_01_C, SDN_01_C, and TZN_01_C and from a study comparing the results of conspicuous observation and simulated client surveys in Benin (see Workbook S8).

The prevalence of health issues

Measuring the prevalence of health issues is complicated by the fact that IMCI algorithms provide several sets of mutually exclusive classifications for diarrhea and fever. All

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

children <5 years old who have diarrhea are to be classified for dehydration, 'persistent diarrhea' and 'dysentery'. Children 2-59 months old who have a fever receive an initial classification focused on malaria or, where malaria is not present, 'possible bacterial infection'. They frequently should be classified for measles and, in a few instances, for dengue as well. Because there are no overlaps among classifications for dehydration, persistent diarrhea, and dysentery, I treat all positive classifications for these conditions as countable health issues. In the case of fever, however, there is a systematic overlap between the classifications for fever per se and those for measles and dengue. Initially, all children presenting with a fever but with no signs of severe febrile disease, malaria, or, where it is relevant, bacterial infection should be classified as 'fever-malaria unlikely' or 'bacterial infection unlikely'. In principle, one or the other of these classifications should be given to all children subsequently found to have measles or dengue, though children suffering from fevers with other etiologies should be classified in this manner as well. As a result, counting classifications for measles and dengue as separate health issues would result in an overcount.

The sick infants and children in the 39 profiles presented with multiple health issues (Figures 2.1 and 2.2). Counting all positive classifications for diarrhea, excluding classifications for measles and dengue, ignoring missing information, and counting 'other problems', where they are noted, as a single concern, the number of issues per patient in the infant populations ranges from a low of 0.8 in KHM_01_I to a high of 2.5 in IND_06_I with a mean of 1.4. Using the same approach, the number of health issues per patient in the child populations ranges from 1.4 in IND_08_C to 3.3 in TZN_02_C, UGA_01_C, and ZWE_01_C with a mean of 2.6.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The patients in the 39 profiles also present with diverse sets of health issues (Figures 2.1 and 2.2). Here I focus on the issues that are used in all the populations in each age group and that are reported in all or nearly all cases. In the infant populations, the prevalence of 'possible bacterial infection' ranges from a low of 0.2 in BOL_01_I, KHM_01_I, and IND_07_I to a high of 0.7 in KEN_02_I. It is the most common symptom in 6 populations. The prevalence of 'feeding problem' ranges from 0.0 in GHN_01_I to 0.9 in IND_05_I. It is the major concern in 5 populations. The prevalence of 'diarrhea' generally is lower, but reaches 0.3 in IND_02_1 and KHM_01_I. In the latter, classifications for 'diarrhea' are slightly more common than for 'possible bacterial infection' and 'feeding problem'.

The most prevalent classifications in the child populations are 'cough and difficult breathing' and 'fever'. The prevalence of the first ranges from 0.4 in BFA_01_C and IND_03_C to 0.9 in BFA_01_C and BGD_04_C. The prevalence of 'fever' ranges from 0.0 in ETH_01_C to 0.9 in BEN_01_C, GMB_01_C, KEN_01_C, KEN_03_C, and VNM_01_C. The prevalence of classifications related to 'diarrhea' ranges from 0.1 in BEN_01_C and ETH_01_C to 0.6 in AFG_01_C, where it is the second most frequent concern. The prevalence of 'ear problem' ranges from 0.00 in BRA_01_C, CHN_01_C, IND_01_C, IND_03_C, KEN_03_C, KHM_01_C, TZN_01_C, TZN_02_C, UGA_01_C, and VNM_01_C to 0.3 in AFG_01_C. The prevalence of classifications for 'malnutrition' is usually quite low, but rises to 0.5 in KEN_01_C. The prevalence of 'anemia' ranges from 0.0 AFG_01_C, AFG_02_C, BGD_03_C, BGD_04_C BWA_01_C, and GMB_01_C to 0.8 in BRA_01_C, where it is the most prevalent symptom.

IMCI tasks and their triggers

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The tasks provided for patient care by the IMCI algorithms fall into three broad groups (Figures 3.1 and 3.2). The spine of a consultation is formed by a series of mandatory tasks. Conditional assessments occur in groups triggered by the result of a mandatory assessment or a previous conditional assessment. Nearly all treatment and counseling tasks occur when all mandatory tasks and conditional assessments have been performed and the resulting classifications are recognized. However, in the Egypt and Morocco algorithms treatment for wheeze is initiated as soon as it is discovered. (For a complete list of tasks and their triggers, see Workbook S7).

Mandatory Tasks. The original generic version of the infant algorithm includes 24 mandatory tasks: a greeting and 4 opening questions, 12 assessments for possible bacterial infection, 1 check for diarrhea, 4 assessments for nutritional status, and checks for immunization status and 'other problems'. The India adaptations provide for 25 mandatory tasks, adding 1 assessment for jaundice, a condition treated in the generic algorithm as an 'other problem'.

The original generic algorithm for children 2-59 months old provides 18 mandatory tasks: the 5 routine opening tasks, 3 assessments for the presence of danger signs, 1 check for each of the 4 'main symptoms', 4 assessments for nutritional status, 1 assessment of immunization status, and a check for other problems. The number of mandatory tasks in subsequent algorithms ranges from a low of 18 in the India version to a high of 24 in the KwaZulu-Natal adaptation. Most of the tasks added in subsequent adaptations have to do with regional or historical differences in important health issues. Throat problems are a concern in Afghanistan, Egypt, and Morocco (4 tasks). HIV is a concern in Botswana, KwaZulu-Natal, and

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Zimbabwe (4 tasks). The 2008 generic, Afghanistan, KwaZulu-Natal, and Morocco algorithms add 'convulsing now' to the list of danger signs. To shorten wait times for very sick children, KwaZulu-Natal requires providers to 'do rapid appraisal on all waiting children' before each new consultation.

Clinic staff should weigh children 2-59 months old and take their temperatures with a thermometer before the consultation with the examining health worker [26]. In the context of this paper, taking the temperature is the most important of these tasks because it takes a minimum of 2 minutes. Practice undoubtedly deviates from these expectations, but the extent of deviation is measured only in health facility surveys for Egypt, Morocco, and Sudan [27, 31, 41]. In these studies, I take the number of times health workers should weigh a child and take her temperature as the proportion of all children for whom that task has not been performed previously by other staff. By this reckoning, the provider involved in the consultation should perform 2 percent of temperature checks and 0.3 percent of weighings in Egypt, 43 percent of temperature checks and 17 percent of weighings in Morocco, and 96 percent of temperature checks and 69 percent of weighings in Sudan. In all other cases, I assume that weighings and temperature checks are performed by the examining health worker and that a thermometer is not used.

Conditional Assessments. The applicable infant algorithms provide three sets of conditional assessments for which the triggers can be tracked. An infant found to have fast breathing should have the breaths in one minute counted a second time. A history of diarrhea triggers five assessments. Providers should perform a 4-minute assessment of breastfeeding for

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

an infant who does not require urgent referral to hospital, but does have difficulty feeding, is breastfed less than 8 times in 24 hours, receives other foods or drinks, and/or is low weight.

The original generic algorithm for children 2-59 months old includes 4 conditional assessments for 'cough or difficult breathing', 6 for 'diarrhea', 6 for 'fever', 3 for 'measles', 4 for 'ear problem', and 3 for feeding issues in children who do not require urgent referral and who are either (a) 'very low weight', (b) anemic, or (c) well-nourished but less than 2 years old.

Subsequent versions made the following changes:

- For a child who is not visibly awake, all add a check for lethargy or unconsciousness.
- For a child with 'cough or difficult breathing', the 2008 generic version and the Afghanistan, Egypt, KwaZulu-Natal, and Morocco adaptations add 'look and listen for wheeze'. The Morocco adaptation also adds 'has the child had recent contact with tuberculosis'.
- For a child with diarrhea, the Kwa-Zulu Natal version adds 'ask what treatment is the mother giving'.
- For a child who presents with a fever, the KwaZulu-Natal version drops the tasks related to measles and the Egypt version drops the redundant 'look for runny nose'. The India, Morocco, and KwaZulu-Natal adaptations add 'look and feel for bulging fontanelle'. The Morocco version also adds 'has the child had recent contact with tuberculosis'. The KwaZulu-Natal algorithm adds 'look for other causes of fever'.
- For a child with measles, the Morocco version drops 'look for clouding of the cornea'.

In addition, the KwaZulu-Natal adaptation calls for 3 assessments to determine the HIV status of children whose status is currently unknown and who have at least 1 of the following: '(a) classification today of pneumonia, persistent diarrhea or not growing well; or (b) episode of

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

persistent diarrhea in last 3 months; or (3) discharge from ear at any time, or (c) mother is HIV positive'.

Assuming that Cambodia and Vietnam use something like the Timor-Leste algorithm for dengue, a fever that has lasted ≥2 days but <7 triggers 7 additional assessments. Since the duration of fever is never reported, I make the conservative assumption that children with no positive classification for dengue did not present with a fever of the required duration and, therefore, did not need these assessments.

Treatment and Counseling. To construct lists of trackable treatment and counseling tasks, I have modified the material found in the original algorithms in three ways. The two infant algorithms contain several tasks triggered by feeding problems that cannot be tracked individually, but for which limited information on indications of use is reported in the observational studies. I replace these tasks with 'counsel mother about feeding issues, thrush, and/or keeping low weight infant warm'.

In the child algorithms that include measles, children with the classification 'measles with eye or mouth complications' have 'pus draining from the eye' and/or 'mouth ulcers'. The algorithms provide treatments for each complication, but the observational studies report only the less specific classification. I deal with this problem by replacing the specific treatments – 'if pus draining from the eye, treat eye infection with tetracycline eye ointment' and 'if mouth ulcers, treat with gentian violet – with 'treat the complications of measles'.

In both infant and child populations, I assume that patients with severe dehydration are referred urgently to hospital rather than given Plan C rehydration locally. If a provider is adequately trained and if the clinic in which she works has the necessary supplies, the preferred

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Plan C option is to administer Ringer's Lactate Solution or normal saline solution for 6 hours, with reassessment and adjustment every 1 or 2 hours [26]. If IV treatment cannot be provided locally or nearby, a second option is to administer ORS solution by mouth, if the infant can drink, or, if the infant is unable to drink and the provider is properly trained, by naso-gastric tube, again for 6 hours with regular reassessment. If the infant is unable to drink, it should be referred urgently to hospital for IV or naso-gastric rehydration. My guess is that the constraints of training, equipment, supplies, and time are such that the option of urgent referral is the first resort far more often than not. In any case, it is impossible to know how, in fact, such patients are treated.

The infant algorithms provide 20 trackable treatment and counseling tasks.

The original generic algorithm provides 26 treatment and counseling tasks that are at least partially trackable for children 2-59 months old. The number of trackable tasks provided by subsequent algorithms ranges from 25 in the Egypt versions to 32 in the KwaZulu-Natal adaptation. Most of the differences relate to the presence or absence of particular diseases. The KwaZulu-Natal adaptation includes 6 tasks related HIV. The Egypt and Morocco adaptations exclude tasks related to malaria. Judging by the portion of the Timor-Leste algorithm concerned with dengue, which I take to apply to Vietnam and Cambodia, caring for patients with a classification of 'dengue hemorrhagic fever' or 'possible DHF' requires no tasks not already provided for fever. Morocco also adds two counseling tasks: 'counsel mother on how to stimulate the psychosocial development of her child' and 'counsel the mother on hygiene'.

There remain several treatment and counseling tasks based, at least in part, on triggers that cannot be tracked in 1 or more of the available studies. The India version of the infant

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

algorithm includes 2 triggers that are not reported: low body temperature and the mother's experience of pain while breastfeeding. For children 2-59 months old, unreported triggers include the presence of wheeze, the presence of cholera in the area, a very high fever, fever for more than a specified number of days, and the presence of cough for more than 30 days. The prevalence of preventive counseling on feeding practices in ZAF_01_C cannot be determined because the study on which it is based does not report the incidence of anemia.

In addition, I make no attempt to track 2 treatment and counseling tasks that have no clear behavioral referent: 'take stock' (in French '*faire bilan*') for Morocco, and 'treat apparent causes of fever' for KwaZulu-Natal, Morocco and Egypt.

Tasks required by patient populations

For infants, the expected total number of trackable tasks per patient ranges from a low of 27.1 in ZAF_01_I to a high of 31.0 in IND_02_I with a mean of 29.5 (Figure 4.1). Three of the tasks expected for infants have a defined minimum duration: a mandatory check of fast breathing requiring at least 1 minute; a conditional check of fast breathing, also requiring at least 1 minute, and a conditional assessment of breastfeeding requiring 4 minutes.

The number of trackable tasks required for children 2-59 months old has a much wider range, from a low of 25.4 in IND_08_C to a high of 45.7 in ZAF_01_C (Figure 4.2). The mean is 36.4. Two tasks expected for children have a minimum duration: a mandatory 'take the temperature with a thermometer' requiring a least 2 minutes and a conditional check of fast breathing requiring at least 1 minute.

Minutes per observed task with no specified minimum duration

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The mean time required by IMCI-trained providers to correctly perform tasks with no minimum duration ranges from 0.2 minutes in TZN_01_C to 1.4 minutes in BENIN_CO and BENIN_SC (Table 1). Two of the studies require no correction for the presence of an observer. In the Benin simulated client survey, the observer was disguised as the patient's caretaker [28]. In BRA_01_C, providers were followed, where possible, for a 5-day working week and no difference was found in the duration of consultations between the first and subsequent days [42]. In the other cases, the minutes required to perform a task with no minimum duration are corrected using values of K_m, K_c, and Q derived from Rowe et al. [28].

In none of these cases are providers doing satisfactory work. Performance rates for mandatory tasks, corrected for observer effects where necessary, range from 0.5 in SDN_01_C to 0.7 in TZN_01_C. Performance rates for conditional tasks are consistently worse, ranging from 0.3 in BRA_01_C to 0.5 in KEN_01_C. Substantial numbers of children are incorrectly and/or incompletely classified. In the populations for which data are available, between 7.3 and 37.4 percent of children are prescribed an unnecessary antibiotic.

Lacking a larger sample or a model case, I take 0.58 minutes, the geometric average of the 7 cases weighted by the number of performance occasions observed, as a reasonable first approximation of the time required to perform the average IMCI task with no specified duration. No great store should be set by this figure, but it probably is not so high that program managers using it would be expecting too little of providers nor so low that they would be expecting providers to perform miracles of speed and efficiency.

Mean expected duration of consultations

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

IMCI consultations require substantial amounts of time (Figures 5.1 and 5.2). For the infant populations, the mean expected duration ranges from 16.7 minutes in ZAF_01_I and IND_01_04 to 18.7 minutes in IND_02_I. The average over all the infant populations is 17.7 minutes. For populations of children 2-59 months old, the range is from a low of 14.9 minutes in IND_08_C to a high of 26.3 minutes in ZAF_02_C, with an average of 21.4.

At the lower end, the various IMCI algorithms provide a floor on the expected duration of consultations that should not be breached. The level of this floor is in large part determined by the provision of mandatory tasks. In the original generic algorithm for infants and in the India version, these tasks are expected to take 14.3 and 14.9 minutes respectively. In the child populations, the mandatory tasks included in the original generic algorithm are expected to take 10.4 minutes. This figure rises to 11.0 minutes in the 2008 generic algorithm, 12.0 in the Sudan version, 13.1 minutes in the Egypt version, 13.3 in the Afghanistan version, 14.0 in the Morocco version, and 13.9 in the KwaZulu-Natal version. The floor also includes a very few treatment and counseling tasks. For example, the original generic algorithm calls for providers to advise all mothers when to return and to counsel them about their own health unless the child requires urgent referral. Depending on the case mix presented by the patient populations and the details of the relevant algorithm, such tasks increase the minimum expected duration of consultations by 0.6 to 2.7 minutes for the infant populations and 0.0 to 5.2 minutes for the child populations.

Finally, consultations with much longer expected durations are not uncommon. In the infant populations, the 98th percentile of expected durations ranges from 20.1 minutes in KEN_02_I, where it is 2.3 minutes longer than the mean, to 25.7 minutes in KEN_02_I, where it is 5.3 minutes longer than the mean. In the child populations, the 98th percentile of expected

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

durations ranges from 21.4 minutes in IND_08_C, where it is 6.5 minutes longer than the mean, to 33.7 minutes in ZAF_02_C, where it is 7.5 minutes longer than the mean. In other words, 1 in 50 consultations is likely to require substantially more time than the average.

DISCUSSION

Few studies have examined the time required for primary care consultations; none have focused on sick child visits in LMICs. This study begins to fill that gap by providing evidence-based estimates of the time needed for initial visits with under-five infants and children at public or not-for-profit facilities in LMICS that have adopted the IMCI clinical algorithm. These figures are based on (1) classification profiles found in, or constructed from, health facility surveys, IMCI impact studies, evaluations of the utility of clinical signs, and similar research; (2) appropriate versions of the IMCI algorithm, and (3) an estimate of the mean time required to perform a task with no pre-defined minimum duration. The last of these is based on a subset of studies of IMCI services that report the health issues presented, task performance rates, and the mean observed duration of consultations.

Limitations

The study has several limitations. First, the information needed to construct classification profiles is incomplete. Among the infant populations, information related to 'persistent diarrhea' and 'dysentery' is missing for KEN_02_I and information related to 'local bacterial infection', 'diarrhea', 'persistent diarrhea' and 'dysentery' is missing for ZAF_01_I. Among the child populations, information is missing for 'persistent diarrhea' and 'dysentery' in BEN_01_C; measles in TZN_01_C and ZAF_02_C; 'ear problem' in BFA_01_C and ZAF_02_C, and 'anemia' in ZAF_02_C. 'Mother's HIV status' is missing for BWA_01_C. Information on the

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

incidence of 'other problems' is incomplete or unavailable for 3 of the 12 infant populations and 25 of the 27 child populations. Figures for immunization status are missing for 10 of the 12 infant profiles and 14 of the 27 child profiles. The size of age groups is missing for 10 of the child populations.

Second, the available studies fail to report the prevalence of several of the conditions that trigger IMCI treatment and counseling. Studies of infant populations in India fail to report low body temperature and the mother's experience of pain while breastfeeding. In the child populations, unreported triggers include, but are not limited to, wheeze, cholera in the area, very high fever, prolonged fever, and prolonged cough.

Third, though the synergistic relationships among nutrition, immunity, and infection are well established [21, 43] and though co-morbidities are a focal concern of IMCI [16], only 1 of the studies on which this paper is based reports patterns of association among IMCI classifications and it excludes classifications for malnutrition and anemia [44]. As a result, I have assumed that classifications and other findings co-occur randomly. This may well not be the case.

Finally, the information required to estimate the mean time required to perform tasks with no minimum duration is available for just 7 populations. Two of these populations, the conspicuous observation and simulated client surveys in Benin [28], from which I derive the coefficients used to adjust for the observer effect, are very small.

Missing information regarding the incidence of classifications and other findings tends to reduce the expected mean duration of consultations. However, except for 'diarrhea' and

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

'anemia' these effects are unlikely to be large. The effects of the third and fourth weaknesses are not known.

Conclusions

IMCI consultations for infants and children presenting for the first time with a new concern require considerably more time than previously recognized. The PHCP suggests that for some patients satisfactory consultation may take as little as 10 minutes [4]. However, the shortest expected consultations in the patient populations examined here exceed 10 minutes by 60.8 to 95.6 percent in the infant and 4.4 to 68.2 percent in the child populations. The WHO suggests that the mean duration of initial consultations with sick children in primary care facilities using the IMCI algorithm is 15 minutes [17]. However, in infant populations, the mean expected duration of consultations exceeds this figure by 18.0 to 38.3 percent. In the child populations, only IND_08_C (mean expected duration 14.9 minutes) conforms to the WHO's expectations. In the other child populations, the mean expected duration of consultations exceeds 15 minutes by 7.8 to 75.1 percent. The expected duration of the 98th percentile of consultations in exceeds 15 minutes by 34.3 to 71.4 percent in the infant populations and by 42.8 to 124.9 in the child populations.

Implications

Estimates of the mean expected duration of IMCI consultations and the steps required to produce them have implications for (a) the methods used to measure provider performance, (b) the facility-level personnel costs associated with operating IMCI facilities, and (c) the quality of care.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Measuring Provider Performance. Survey instruments designed to measure the quality of care delivered to sick children should be cognizant of the fact that the standards against which the performance of IMCI-trained providers is measured are not fixed. This is, in part, a function of the design of IMCI algorithms. In all versions of the algorithm, the set of tasks to be performed during follow-up visits overlaps with but is different from that to be performed during initial visits. Moreover, to the original generic version of the algorithm for initial visits, subsequent adaptations added as many as 6 mandatory tasks; moved into the foreground concerns such as jaundice, throat problems, HIV status, dengue, and wheeze; and revised the triggers for many treatment and counseling tasks (Workbook S7).

Expected performance rates for conditional assessments and for treatment and counseling tasks also vary with the classification profiles of patient populations. Among the child populations examined in this study, for example, ETH_01_C and KEN_01_C both employ the original generic IMCI algorithm. However, providers in ETH_01_C are expected to perform, on average, just 31.2 tasks, while providers in KHM_01_C are expected to perform 38.7, 24.0 percent more. Something similar is seen in the infant populations, as well, though in this age group the differences are not so large. Thus, IND_06_I and IND_09_I BOTH use the India adaptation of the algorithm. Nevertheless, providers working in IND_06_I are expected to perform 4.0 percent more tasks than those working in IND_09_I, 31.1 versus 29.9.

Survey instruments designed to assess the quality of care also should recognize that IMCI-trained providers not infrequently classify their patients erroneously. A systematic review of the effects of IMCI training found that 10 to 67 percent of children seen by IMCI-trained providers are given at least 1 incorrect classification [45].

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The Health Facility Survey (HFS) [17], the first instrument designed to measure the effects of IMCI training at the national level, deals with variation in the composition of patient populations by kind of visit and age by focusing on initial visits and by reporting separately on infants <2 months old and children 2-59 months old. It deals with differences in patient population classification profiles and potential errors in provider classifications by defining key indicators of provider performance as proportions. The numerator of each indicator is the number of times task *n* is performed when indicated. The denominator is the number of children with validated classification(s) for which task *n* is indicated. Validated classifications are those "made by an IMCI-trained expert clinician after re-examining the child" after the standard visit with the IMCI-trained provider [17].

Unfortunately, the child health component of DHS Program's now widely used Service Provision Assessment (SPA) [46], makes it difficult or impossible to calculate such proportions. SPA samples include under-five children presenting for the first time with new issues and for follow-up care. Though the observation protocol calls for collection of the age of each patient and the purpose of the visit, final SPA reports lump all of this together. The SPA also fails to provide validated or gold standard or validated classifications, presumably because the expert reassessments on which they are based are time consuming and expensive. Instead, the observation protocol for sick child services calls for the researcher to "ask the provider to tell" her "the diagnosis for the sick child" [47]. As a result, the SPA provides satisfactory counts of tasks performed but, except for mandatory tasks, nothing on the number of children for whom tasks are indicated. Researchers who wish to determine the proportion of children for whom

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

mandatory task *x* is performed will have to perform a patient-by-patient analysis of the relevant SPA raw data to separate out those to whom the mandate applies.

High Personnel Costs. Estimates of the mean expected duration of IMCI consultations also have implications for the facility-level personnel costs associated with the IMCI program. The issue here revolves around (a) the time required to perform a sick-child consultation and (b) how best to measure that time. The creators of the IMCI algorithm recognized that it was likely to increase the duration of consultations and that primary care facilities, therefore, might need to employ more health care providers. However, it was hoped that the increases would be modest and that they might be offset by reductions in providers' unused capacity [16, 48]. The findings of this study, and the steps required to achieve them, suggest that increases in the duration of consultations due to the IMCI algorithm likely are substantial. They also suggest that the observed duration of consultations alone is an unsatisfactory measure of the expected duration.

Three studies have attempted to measure of the effects of IMCI on the duration of consultations. Each is closely linked to and comparable with 1 or more of the classification profiles used in this paper. The first study [49] was carried out in Matlab, Bangladesh, the same rural subdistrict as the 3 Bangladesh IMCI evaluation for which I have constructed classification profiles (Workbook S1) and just a year before BGD_02_C, the first in the series. Studies in Brazil [42] and Tanzania [50] are the sources for the data used to construct BRA_01_C and TZN_01_C. In Bangladesh, visits with paramedics and community health workers using a questionnaire based on the IMCI protocol were, on average, well over twice as long as those in which providers used their pre-IMCI practices, 18.8 minutes versus roughly 8 minutes. In the Brazil and

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Tanzania studies, the effect of using the IMCI algorithm was found to be much smaller. In Brazil, consultations with IMCI providers initially appeared to be 71.5 percent longer than those with non-IMCI providers, 11.08 versus 6.46 minutes. However, after controlling for other determinants of consultation length such as time of day, IMCI visits were just 29.2 percent longer [42]. In Tanzania, the average consultation with an IMCI-trained provider was 30.2 percent longer than those in comparison facilities, 8.2 minutes as opposed to 6.3 [50].

The results of the Bangladesh study appear to be reliable. Unfortunately, the study was not observational. Instead, community health workers and paramedics were asked to time themselves with digital watches while managing sick children with the IMCI-based questionnaire. At meetings, providers also were asked to come to consensus agreements on (a) how long it would take to provide "feedings counseling and other general advice not directly related to" the illnesses presented and (b) how long it took them to perform consultations using pre-IMCI procedures. Aware that providers might fail to perform some of the required IMCI tasks, the Matlab researchers used bi-weekly meetings to "ensure" that participating providers performed all the IMCI tasks even when they appeared "unnecessary or irrelevant for a case" [49]. This device appears to have worked reasonably well. The figure of 18.8 minutes for the mean duration of consultations falls well within 2 standard deviations of the mean expected duration of consultations for the 3 Bangladesh classification profiles presented here which range from 21.2 to 21.4 minutes (Workbook S9). The implied time to perform a task with no minimum duration in Matlab also is consistent with my estimate of 0.58 minutes. If, as in BGD_02_C, the Matlab providers were expected to perform 0.7 checks for fast breathing and 36.3 tasks with no minimum duration, and if, which seems unlikely, they performed all required tasks, it would take

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

them on average 0.5 minutes to perform each task with no minimum duration. If they performed 90 percent of expected tasks, which is still a very high level, it would take 0.57 minutes to perform each task with no minimum duration, very close, indeed, to the estimate I provide in the previous section.

On the other hand, the findings of the Brazil and Tanzania studies appear to be artifacts of the methods used to produce them. The results presented here suggest that If providers performed all required tasks and if, in Tanzania, they took what appears a plausible amount of time to perform tasks with no minimum duration then IMCI consultations would take, on average, 20.1 minutes in Brazil and 20.6 minutes in Tanzania. The Brazil and Tanzania studies directly observed sick-child visits and provided gold standard classifications. However, when it comes to the time required for IMCI consultations, they take the observed duration of consultations as an adequate indicator of expected duration, ignoring their own data on task performance rates. The Tanzania study also ignores the likely effects on performance rates of the presence of observers. The task performance rates reported in these and other studies of IMCI services make clear that IMCI-trained providers routinely fail to perform many required tasks (Workbook S8). In Brazil, where the results do not have to be adjusted for the presence of an observer, IMCI-trained providers performed, on average, 84 percent of 10 mandatory tasks and just 48 percent of 5 conditional tasks. In Tanzania, where the results do have to be adjusted for observer effects, IMCI-trained providers performed, on average, 91 percent of 7 mandatory tasks and 72 percent of 6 mandatory tasks [51]. After adjustment, these figures fall to 69 percent for mandatory tasks and 45 percent for conditional tasks.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

In short, in the absence of adjustments for task performance rates, the observed duration of IMCI consultations is very likely a measure of what might be called budget or cut-rate IMCI services in which patients get the short end of the stick.

Poor Provider Performance. Evidence-based estimates of the expected duration of IMCI consultations indicate that the short end of the stick is precisely what the bulk of sick children get. They confirm, yet again, that the performance of health care providers in resource-poor settings is unsatisfactory [35, 52].

It is well established that first visits with IMCI providers are "brief" [35]. A review of Service Provision Assessments (SPAs) in Haiti, Kenya, Malawi, Namibia, Nepal, Rwanda, Senegal, Tanzania, and Uganda found that, excluding "each provider's first observed consultation to minimize the Hawthorne effect" as well as "the longest 1 percent of visits", the median duration of visits ranged from 6 minutes in Malawi, Nepal, Rwanda, and Uganda to 12 minutes in Tanzania. Across the SPA studies, "82 percent of visits were shorter than 15 minutes" [35]. Since SPA reports do not distinguish between initial and follow-up visits, and since follow-up visits undoubtedly are shorter, these findings may underestimate the time spent on initial visits. However, they are consistent with a recent study of rural IMCI services in Cambodia, Guatemala, Kenya, and Zambia. This study found that the mean observed duration of consultations ranged from 8.2 minutes in Zambia to 12.6 minutes in Guatemala [53].

Measured against the estimates of the expected duration of IMCI consultations presented here, these observed visits are, in fact, brutally brief. The bulk of consultations do not last long enough for health workers to perform even the minimum of mandatory tasks, let alone the full range of conditional assessments, treatments, and counseling messages.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Acknowledgements

I am grateful to Dr. Alexander K. Rowe (The *Global Fund* to Fight AIDS, Tuberculosis, and Malaria) for helping me adapt his findings on conspicuous observation and simulated client surveys in Benin [28] to the aims of this study.

Research Ethics

No IRB review was required. All the data used comes from published studies or public online sources. None can be used to identify individuals.

Funding

None

Conflicts of Interest

None

Supporting Information (DOI: 10.5281/zenodo.8436402)

Data S1	Report of the IMCI Health Facility Survey in Botswana, 2007-2008
Data S2	Report of the IMCI Health Facility Survey in Tanzania, 2003
Text S1	Notes on the Construction of Classification Profiles
Text S2	Notes on the Identification of IMCI Tasks
Text S3	Notes on the Estimation of Minutes per Task with No Minimum Duration
Workbook S1	Studies Used
Workbook S2	Classification Profiles for Infants
Workbook S3	Classification Profiles for Children in Upper Middle-Income Countries
Workbook S4	Classification Profiles for Children in Lower Middle-Income Countries, Part I
Workbook S5	Classification Profiles for Children in Lower Middle-Income Countries, Part II

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Workbook S6 Classification Profiles for Children in Low Income Countries

Workbook S7 IMCI Clinical Tasks

Workbook S8 Minutes per Tasks with No Minimum Duration

Workbook S9 Summary Results

References

1. Byrne P, Long B. Doctors Talking to Patients: A Study of the Verbal Behavior of General Practitioners Consulting in their Surgeries. London: H. M. Stationary Office; 1976.
2. Fiscella K, Epstein R. So much to do, so little time: care for the socially disadvantage and the 15-minute visit. Archives of Internal Medicine. 2008;168(17):1843-52.
3. Royal College of General Practitioners. The 2022 GP: A vision for General Practice in the future NHS 2013 [24 June 2022]. Available from: https://www.pulsetoday.co.uk/news/uncategorised/in-full-the-2022-gp-a-vision-for-general-practice-in-the-future-nhs/.
4. Primary Health Care Performance Initiative. Primary Health Care Vital Signs Profiles: Detailed Methodology Note: Primary Health Care Performance Initiative; 2018 [24 July 2021]. Available from: https://improvingphc.org/sites/default/files/2018.10.22%20VSP%20Methodology%20Note_0.pdf.
5. Reyburn, Mwakasungula E, Chonya S, Mtei F, Bygbjerg I, Poulsen A, et al. Clinical assessment and treatment in paediatric wards in the north-east of the United Republic of Tanzania. Bulletin of the World Health Organization. 2008;86:132-9.
6. Hughes D. Consultation length and outcome in two group general practices. British Journal of General Practice. 1983;33(248):143-7.
7. Roland M, Bartholomew J, Courtenay M, Morris R, Morrell D. The “five minute” consultation: effect of time constraint on clinical content and patient satisfaction. British Medical Journal. 1986;292:870-3.
8. Howie J, Porter A, Forbes J. Quality and the use of time in general practice: widening the discussion. British Medical Journal. 1989;298:1008-110.
9. Howie J, Porter A, Porter D, Hopton J. Long to short consultation ratio: a proxy measure of quality of care for general practice. British Journal of General Practice. 1991;41:48-54.
10. Wilson A. Consultation length in general practice: a review. British Journal of General Practice. 1991;41(119-122).
11. Freeman G, Horder J, Howie J, Pali Hungin A, Hill A, Shah N, et al. Evolving general practice consultation in Britain: issues of length and context. British Medical Journal. 2002;324:880-2.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

12. Wilson A, Childs S. The relationship between consultation length, process and outcomes in general practice: a systematic review. British Journal of General Practice. 2002;52(485):1012-20.
13. Deveugele M, Derese A, van den Brink-Muinen A, Bensing J, De Maeseneer J. Consultation length in general practice: cross sectional study in six European countries. British Medical Journal. 2002;235(7362).
14. Dugdale D, Epstein E, Pantilat S. Time and the patient–physician relationship. Journal of General Internal Medicine. 1999;14:S34-S40.
15. Outomuro D, Actis A. Estimacíon del tiempo de consulta ambulatoria en clínica médic. Revista Médica de Chile. 2013;141:361-6.
16. Tulloch J. The integrated approach to child health in developing countries. The Lancet. 1999;354:16-20.
17. World Health Organization. Health Facility Survey: Tool to Evaluate the Quality of Care Delivered to Sick Children Attending Outpatient Facilities Geneva, Switzerland: World Health Organization; 2003. Available from: https://apps.who.int/iris/handle/10665/42643.
18. Druss B, Mechanic D. Should visit length be used as a quality indicator in primary care. The Lancet. 2003;361(9364):1148.
19. Gove S. Integrated management of childhood illness by outpatient health workers: technical basis and overview. Bulletin of the World Health Organization. 1997;75:7-24.
20. Boschi-Pinto C, Labadie G, Dilip T, Oliphant N, Dalglish S, Aboubaer S, et al. Global implementation survey of Integrated Management of Childhood Illness (IMCI): 20 years on. BMJ Open [Internet]. 2018; 8.
21. Scrimshaw NS, SanGiovanni JP. Synergism of nutrition, infection, and immunity: an overview. American Journal of Clinical Nutrition. 1997;66(2):46S-77S.
22. Black R, Morris S, Bryce J. Where and why are 10 million children dying every year. The Lancet. 2003;361(9376):2226-34.
23. Victora C, Adam T, Bryce T, Evans D. Integrated management of the sick child. In: Jamison DT, J G Breman, A R Measham, G Alleyne, M Claeson, D B Evans, P Jha, A Mills and P Musgrove, editor. Disease Control Priorites in Developing Countries. 2nd ed. Washington, DC: World Bank and Oxford University Press; 2006.
24. Rowe AK, Hirnschall G, Lambrechts T, Bryce J. Linking the Integrated Management of Childhood Illness (IMCI) and health information system (HIS) classifications: issues and options. Bulletin of the World Health Organization. 1999;77(12):988-95.
25. Ministry of Health and Family Welfare-India. Students' Handbook for Integrated Management of Neonatal and Childhood Illness New Delhi: Government of India; 2003. Available from: https://medbox.org/document/india-students-handbook-for-imnci-integrated-management-of-neonatal-and-childhood-illness#GO.
26. World Health Organization. Handbook: IMCI Integrated Management of Childhood Illness Geneva: World Health Organization; 2005. Available from: https://apps.who.int/iris/handle/10665/42939.
27. Ministry of Health and Population-Egypt. Health Facility Survey on Outpatient Child Care (IMCI) Services, Egypt, March 2002 Cairo: World Health Organization; 2003. Available from: http://applications.emro.who.int/dsaf/EMRPUB_2002_EN_1370.pdf.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

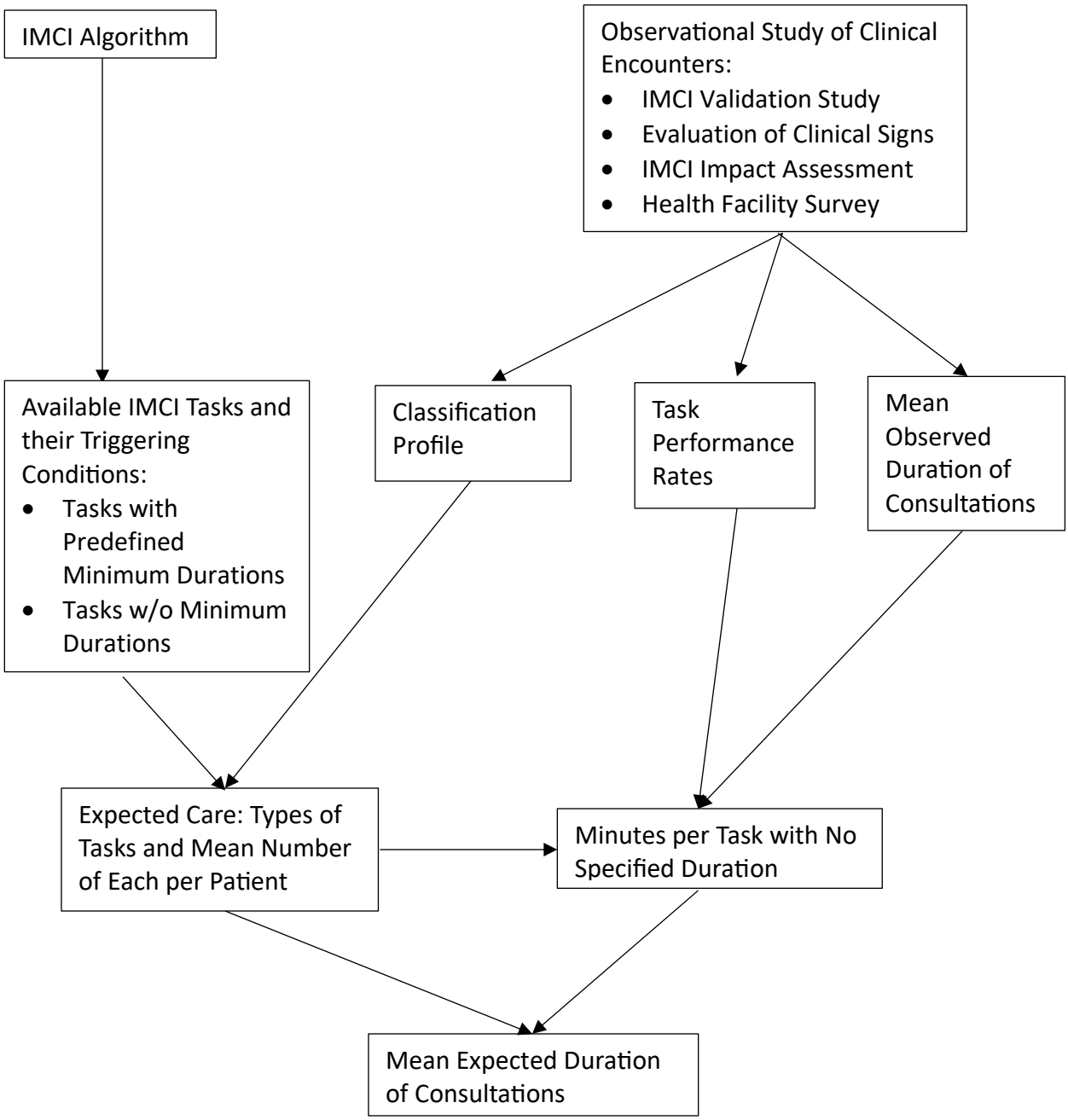
28. Rowe AK, Onikpo F, Lama M, Deming MS. Evaluating health worker performance in Benin using the simulated client method with real children. Implementation Science. 2012;7(1):1-11.
29. Leonard K, Masatu MC. Outpatient process quality evaluation and the Hawthorne effect. Social Science & Medicine. 2006;63:2330-40.
30. Das J, Sohnesen TP. Variations in doctor effort: evidence from Paraguay. Health Affairs. 2007;26(3):w324-w37.
31. Ministry of Health-Morocco. IMCI Health Facility Survey, Morocco, October-December 2007 Cairo: World Health Organization; 2009. Available from: http://applications.emro.who.int/dsaf/dsa973.pdf.
32. Bernasconi A, Crabbé F, Adedeji AM, Bello A, Schmitz T, Landi M, et al. Results from one-year use of an electronic clinical decision support system in a post-conflict context: an implementation research. PloS One. 2019;14:e0225634.
33. Mustafa L. A hospital based prospective observational cross sectional study from India to evaluate the role of integrated management of childhood illness (IMCI) in under five childhood illnesses (6 months to 5 years). Journal of Pediatrics and Neonatal Care. 2016;4(6):151-9.
34. Mæstad O, Torsvik G, Aakvik A. Overworked? On the relationship between workload and health worker performance. Journal of health economics. 2010;29(5):686-98.
35. Kruk ME, Gage AD, Mbaruku GM, Leslie HH. Content of care in 15,000 sick child consultations in nine lower-income countries. Health Services Research. 2018;53(4):2084-98.
36. Boulanger LL, Lee LA, Odhacha A. Treatment in Kenyan rural health facilities: projected drug costs using the WHO-UNICEF integrated management of childhood illness (IMCI) guidelines. Bulletin of the World Health Organization. 1999;77(10):852-8.
37. Mitchell M, Msellem D, Lesh N. Using electronic technology to improve clinical care: results from a before-after cluster trial to evaluate assessment and classification of of sick children according to Integrated Management of Childhood Illness (IMCI) protocol in Tanzania. BMC Medical Informatics and Decision Making [Internet]. 2013; 13(95). Available from: https://rdcu.be/cQtWb.
38. Simoes E, Desta T, Tessema T, Gerbresellassie T, Dagnew M, Gove S. Performance of health workers after training in integrated management of childhood illness in Gondar, Ethiopia. Bulletin of the World Health Organization. 1997;75(Suppl 1):43.
39. Ministry of Health-Zambia. Zambia 2018 National Integrated Management of Newborn and Childhood Illnesses Health Facility Survey Report Lusaka, Zambia: Ministry of Health-Zambia; 2018. Available from: https://www.afro.who.int/publications/zambia-2018-national-integrated-management-newborn-and-childhood-illnesses-health.
40. Jibo AM. Assessment of the Effect of Implementation of the Integrated Management of Childhood Illnesses (IMCI) Approach on Childcare in Kano State Nigeria Zaria, Nigeria: Ahmadu Bello University Teaching Hospital; 2010. Available from: Available at https://dissertation.npmcn.edu.ng/index.php/FMCPH/article/view/2475.
41. Ministry of Health-Sudan. IMCI Health Facility Survey, Sudan, March-April 2003 Cairo: World Health Organization; 2004. Available from: http://applications.emro.who.int/dsaf/EMRPUB_2003_EN_1369.pdf.
42. Adam T, Amorim DG, Edwards SJ, Amaral J, Evans DB. Capacity constraints to the adoption of new interventions: consultation time and the Integrated Management of Childhood Illness in Brazil. Health policy and planning. 2005;20(suppl 1):i49-i57.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

43. Caulfield LE, Richard SA, Black RE. Undernutrition as an underlying cause of malaria morbidity and mortality in children less than five years old. American Journal of Tropical Medicine and Hygiene. 2004;71:55-63.
44. Kolstad PR, Burnham G, Kalter HD, Kenya-Mugisha N, Black RE. Potential implications of the integrated management of childhood illness (IMCI) for hospital referral and pharmaceutical usage in western Uganda. Tropical Medicine & International Health. 1998;3(9):691-9.
45. Nguyen DTK, Leung KK, McIntyre L, Ghali WA, Sauve R. Does Integrated Management of Childhood Illness (IMCI) training improve the skills of health workers? A systematic review and meta-analysis. PLoS ONE. 2013;8(6):e66030.
46. USAID. Service Provision Assessment [29 September 2022]. Available from: https://dhsprogram.com/methodology/Survey-Types/SPA.cfm.
47. Program TD. SPA Questionnaires 2022 English [30 September 2022]. Available from: https://dhsprogram.com/publications/publication-SPAQ8-SPA-Questionnaires-and-Manuals.cfm.
48. Adam T, Bishai D, Khan M, Evans D. Methods for the costing component of the Multi-Country Evaluation of IMCI Geneva: World Health Organization; 2004. Available from: http://www.who.int/imci-mce.
49. Khan MM, Ahmed S, Saha KK. Implementing IMCI in a developing country: estimating the need for additional health workers. Human Resources for Health Development. 2000;4:73-82.
50. Adam T, Manzi F, Schellenberg JA, Mgalula L, de Savigny D, Evans DB. Does the Integrated Management of Childhood Illness cost more than routine care? Results from the United Republic of Tanzania. Bulletin of the World Health Organization. 2005;83(5):369-77.
51. Schellenberg JA, Bryce J, de Savigny D, Lambrechts T, Mbuya C, Mgalula L, et al. The effect of Integrated Management of Childhood Illness on observed quality of care of under-fives in rural Tanzania. Health policy and planning. 2004;19(1):1-10.
52. Rowe AK, de Savigny D, Lanata C, Victora CG. How can we achieve and maintain high-quality performance of health workers in low-resource settings. The Lancet. 2005;366:1026-35.
53. Edward A, Dam K, Chege J, Ghee AE, Zare H, Chhorvann C. Measuring pediatric quality of care in rural clinics — a multi-country assessment — Cambodia, Guatemala, Zambia, and Kenya. International Journal for Quality in Health Care. 2016;24:586-93.

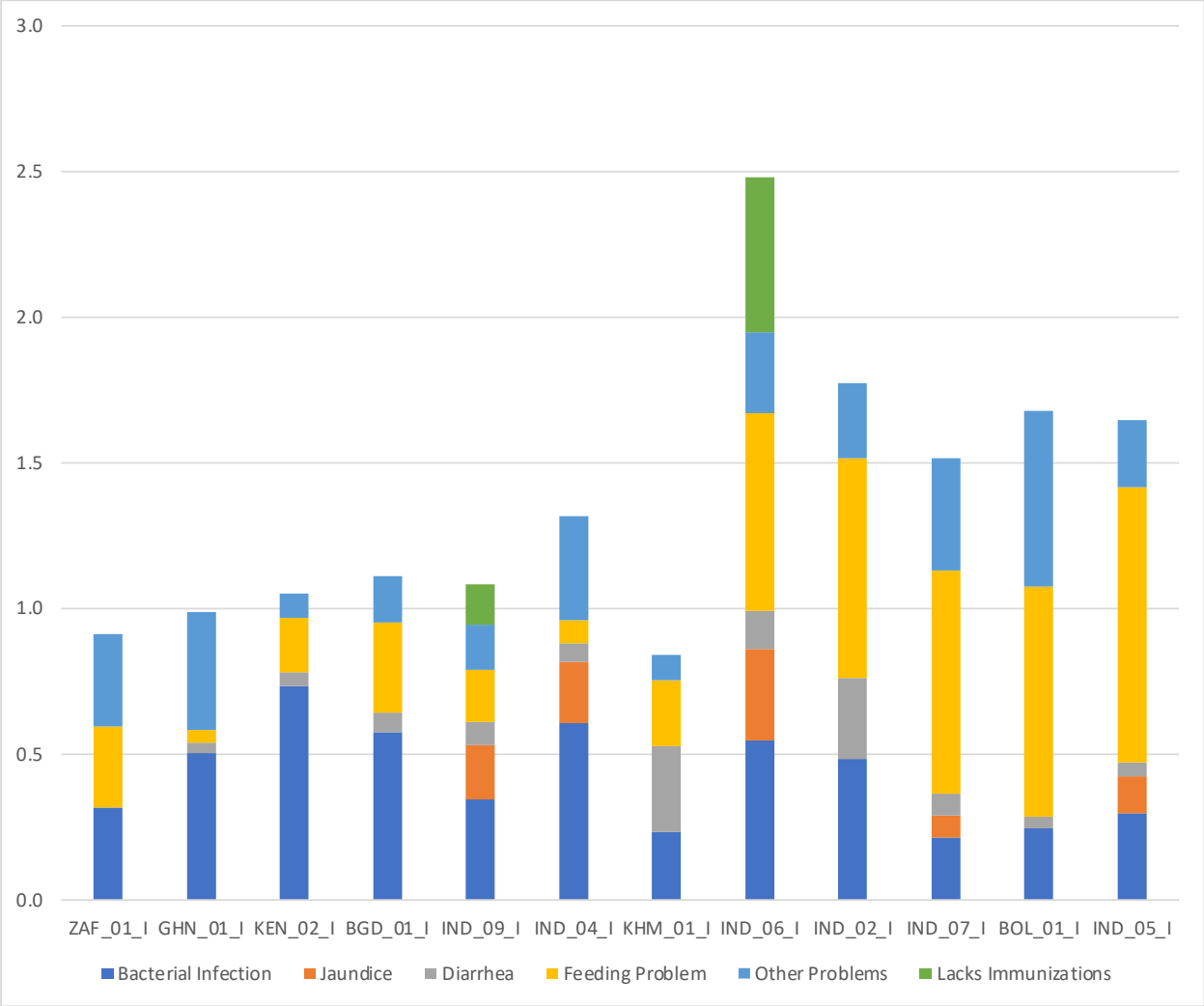
medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 1: Analytical Framework



medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 2.1 Prevalence of Health Issues in Populations of Infants <2 Months Old (Populations Sorted by Expected Mean Duration of Consultation)

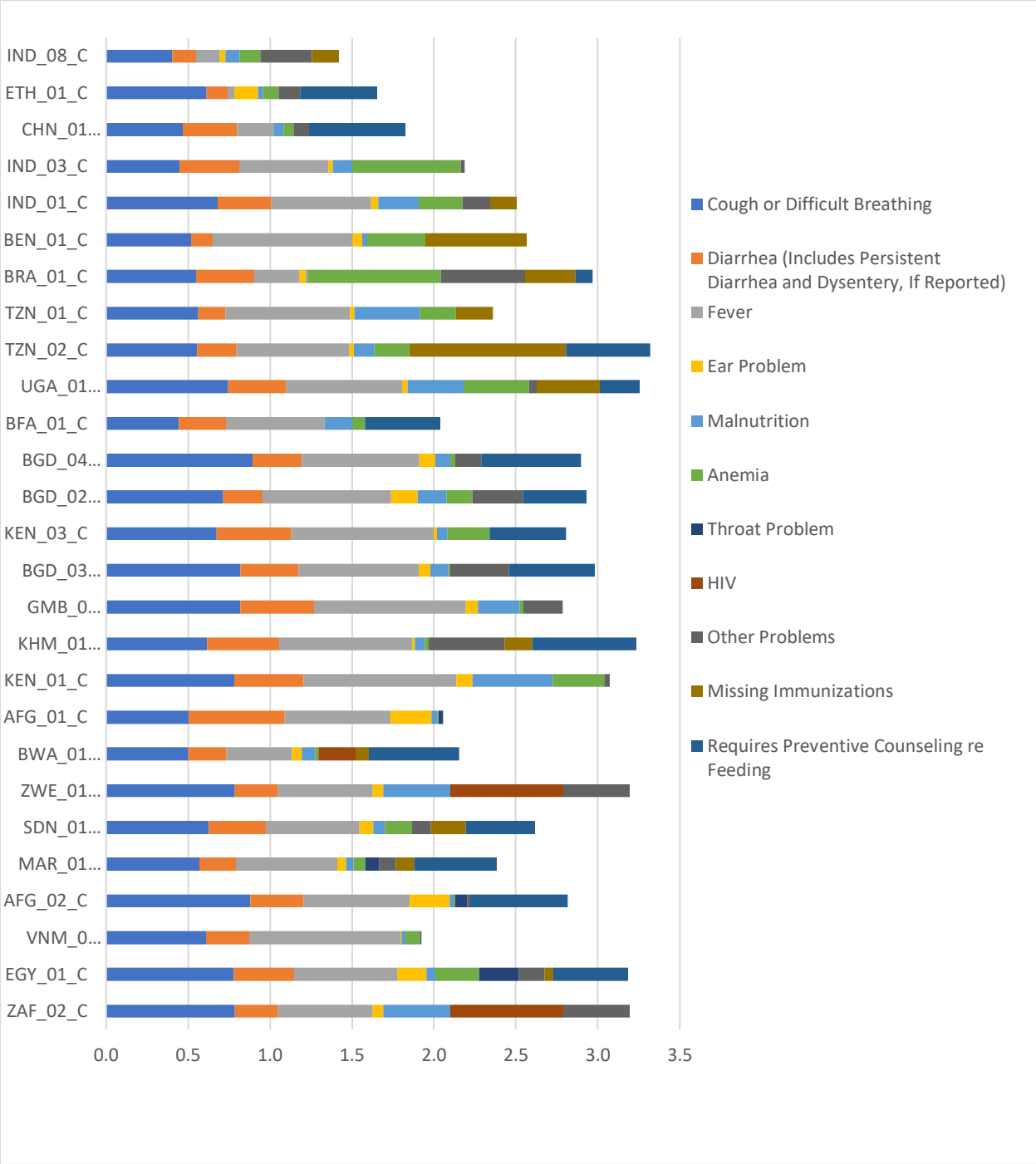


Source: Workbook S9 (Infant Worksheet)

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 2.2 Prevalence of Health Issues in Populations of Children 2-59 Months Old

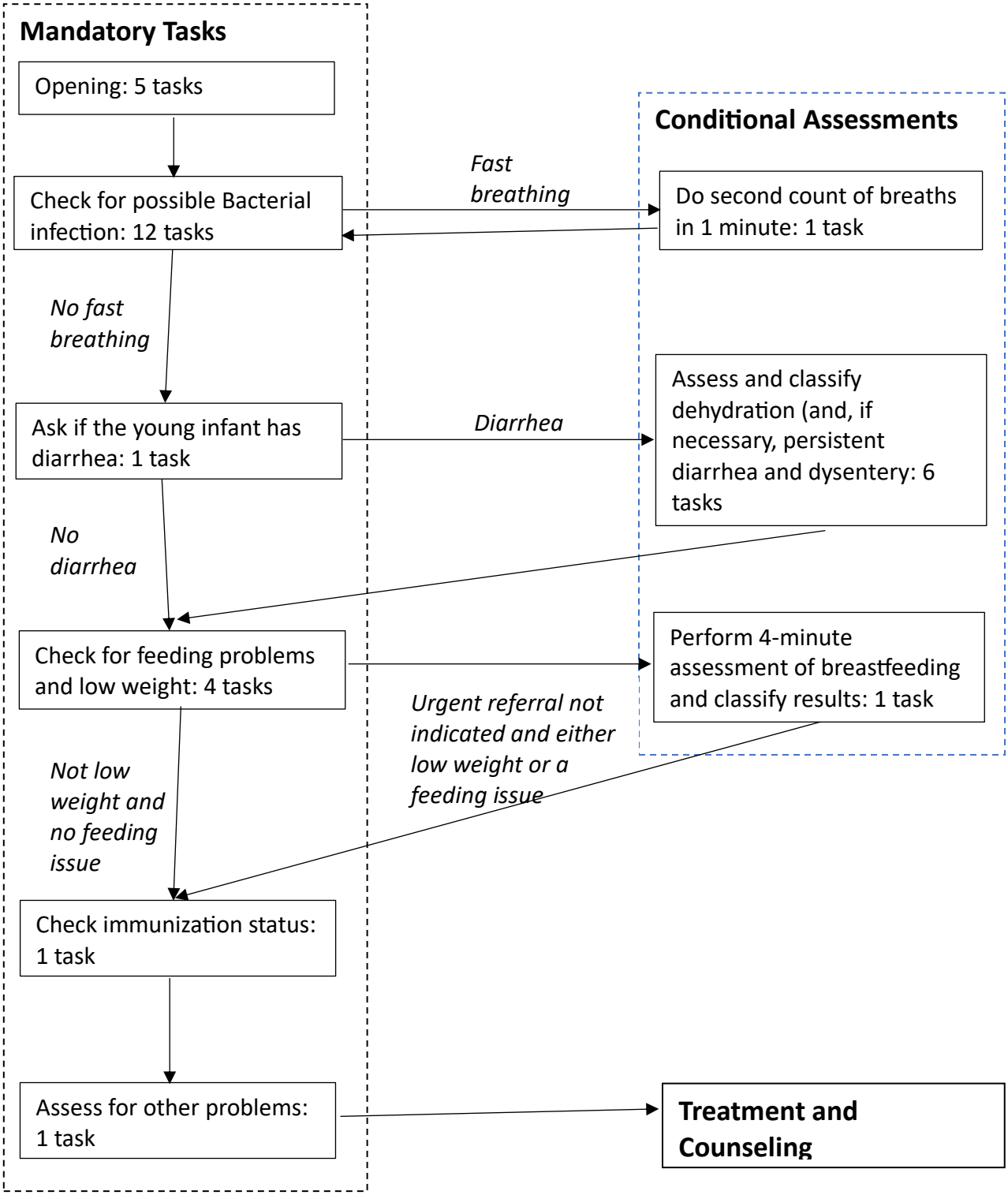
(Populations Sorted by Expected Mean Duration of Consultation)



Source: Workbook S9 (Child Worksheet)

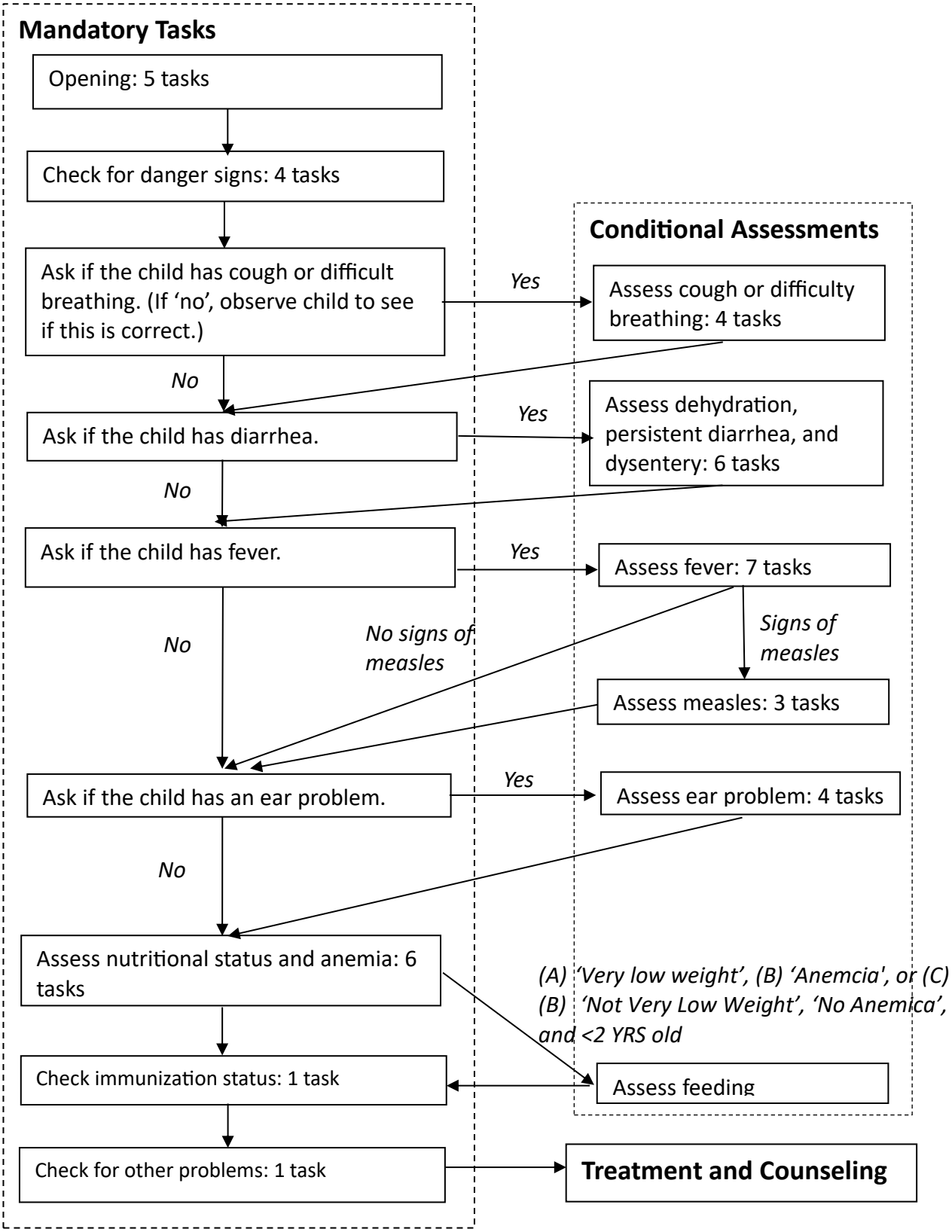
medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 3.1 Outline of Initial IMCI Visits for Infants <2 Months Old Using Generic Algorithm



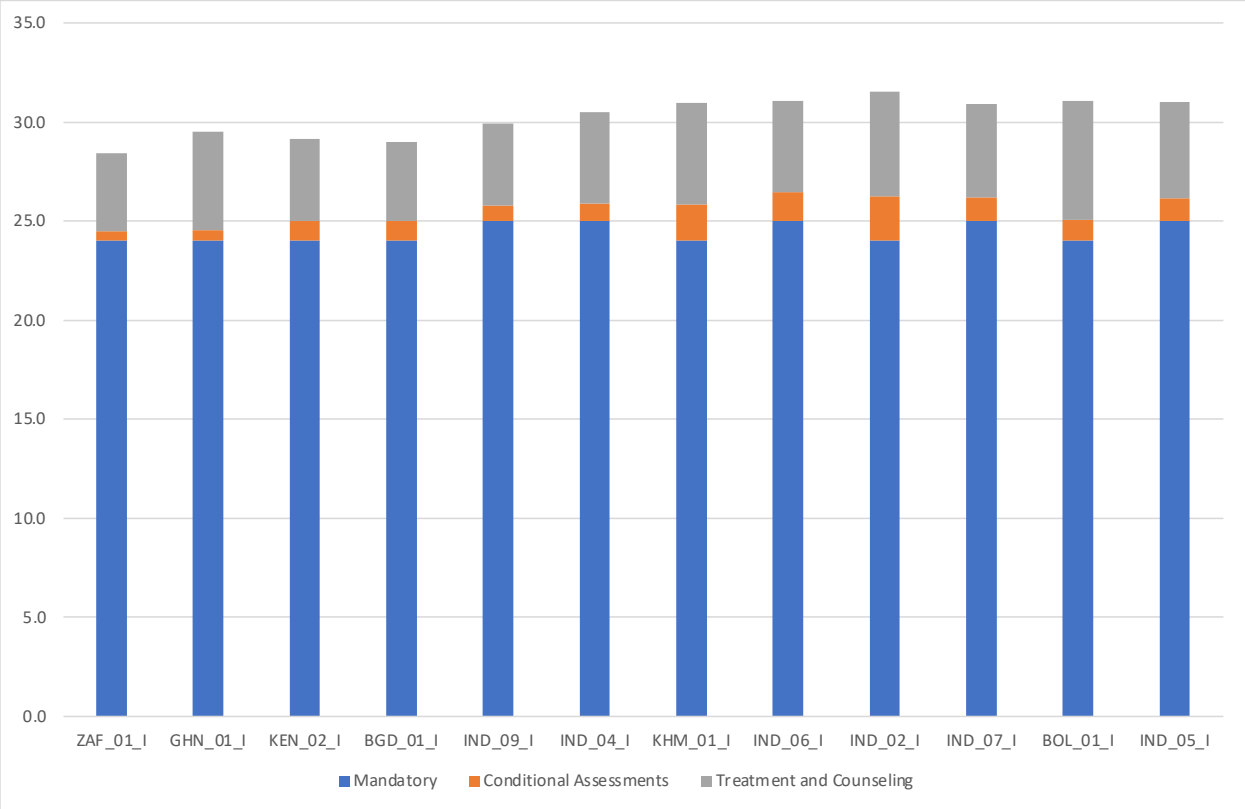
medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 3.2 Outline of Initial IMCI Visits for Children 2-59 Months Old



medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

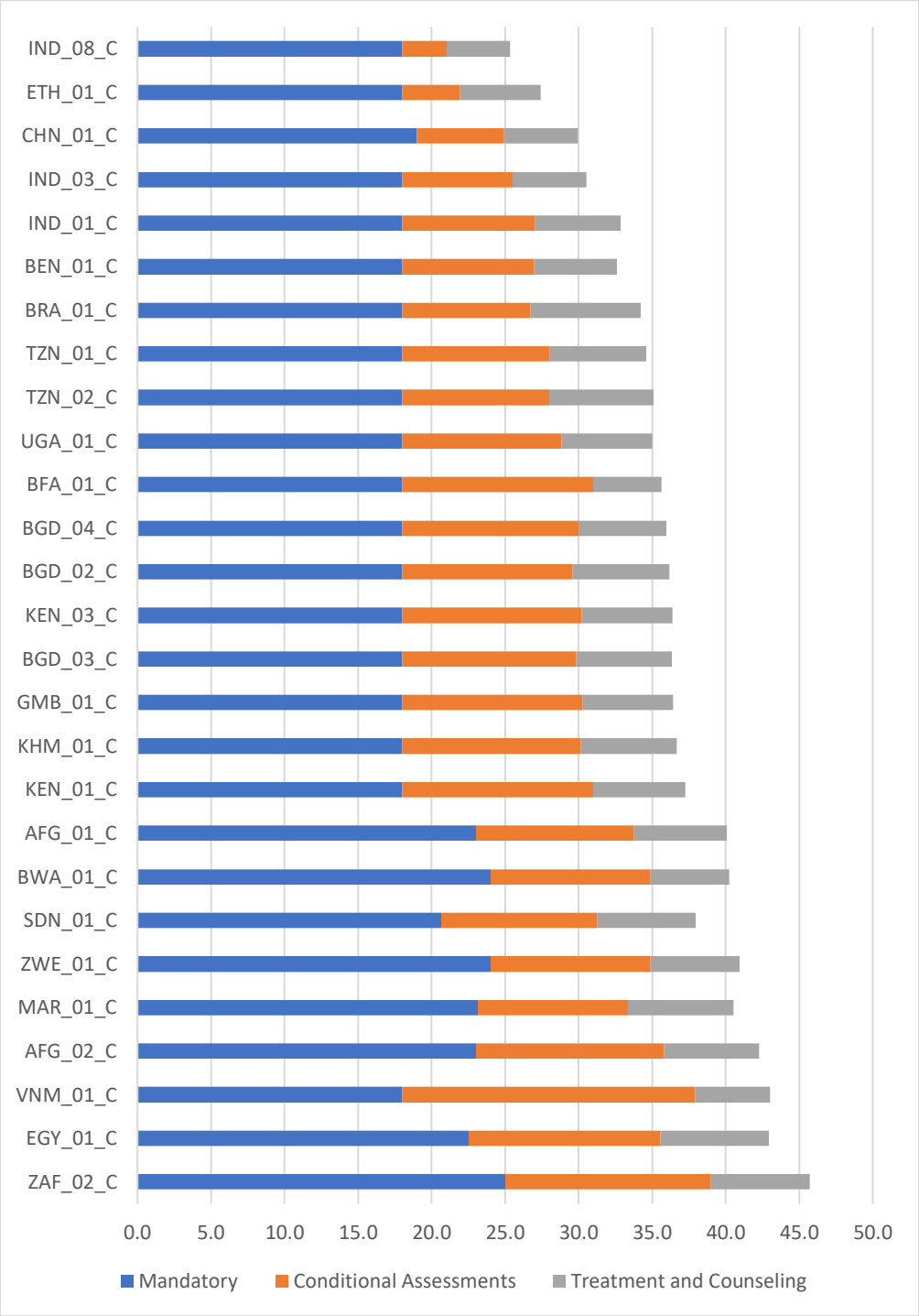
Figure 4.1 Mean Number of IMCI Tasks Required for Populations of Infants <2 Months Old (Populations Sorted by Expected Mean Duration of Consultation)



Source: Workbook S9 (infant Worksheet)

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

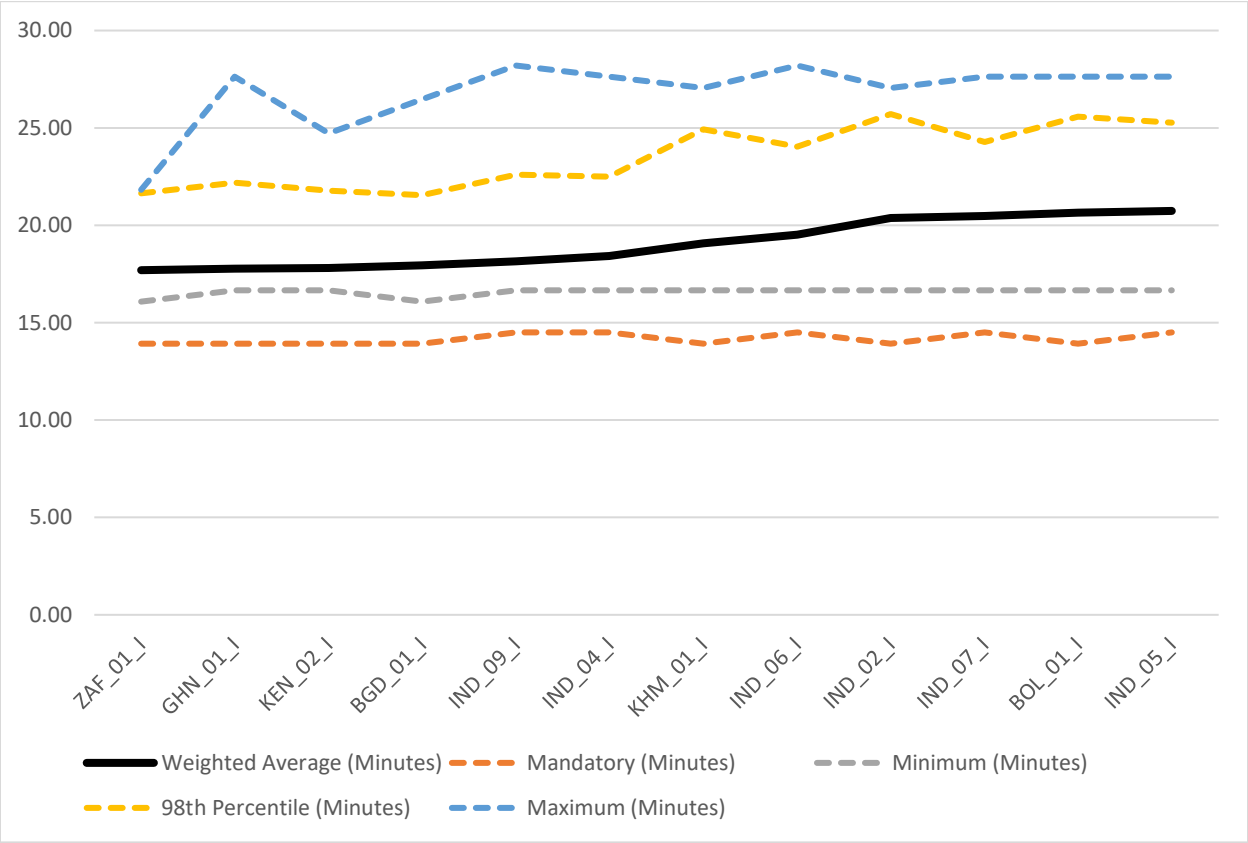
Figure 4.2 IMCI Tasks Required for Children 2-59 Months Old (Populations Sorted by Expected Mean Duration of Consultation



Source: Workbook S9 (Child Worksheet)

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

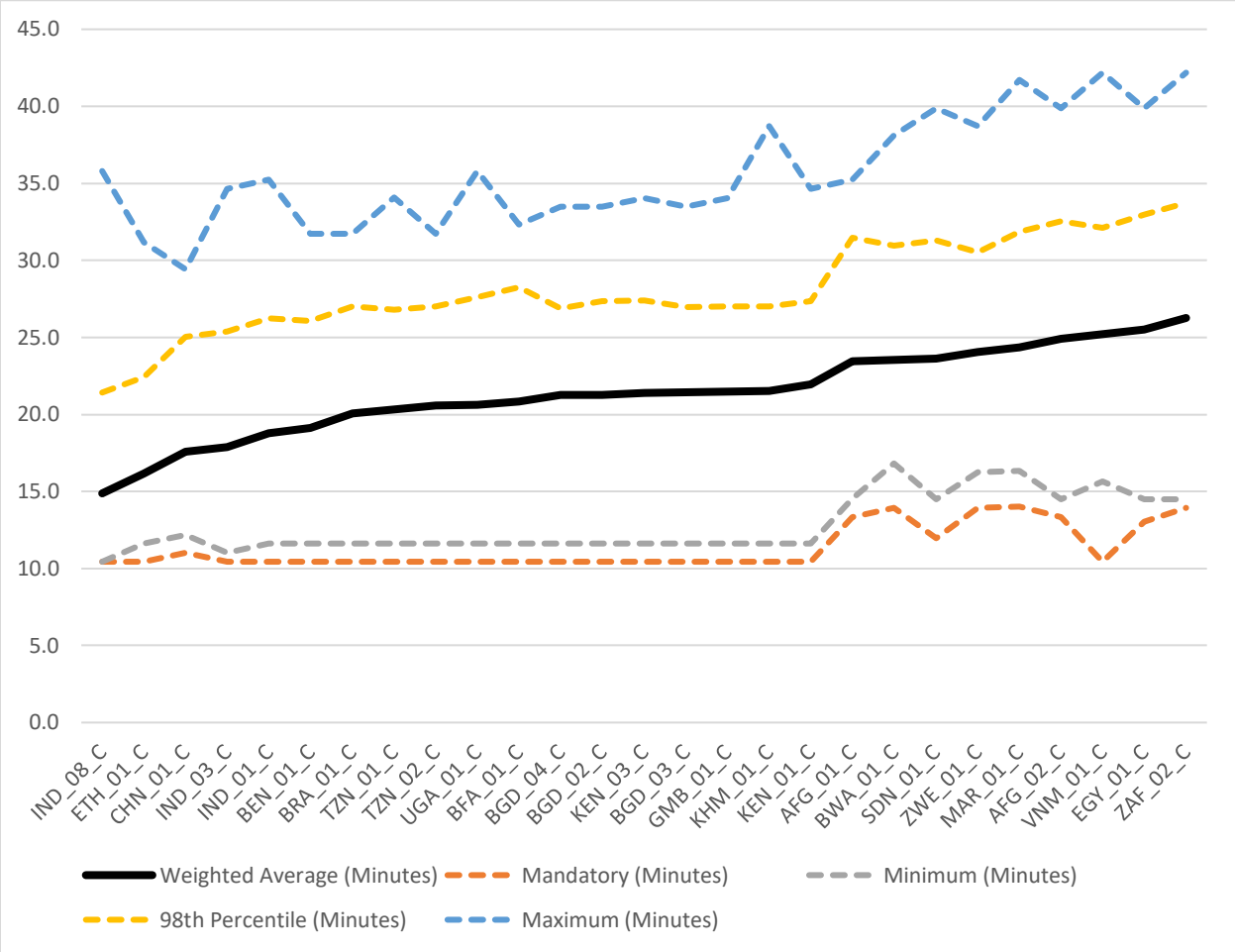
Figure 5.1 Expected Durations of IMCI Consultations with Infants <2 Months Old in Minutes (Populations Sorted by Expected Mean Duration of Consultation)



Source: Workbook S9 (Infant Worksheet)

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 5.2 Expected Durations of IMCI Consultations with Children 2-59 Months Old in Minutes (Populations Sorted by Expected Mean Duration of Consultation)



Source: Workbook S9 (Child Worksheet)

medRxiv preprint doi: https://doi.org/10.1101/2023.06.03.23290926; this version posted October 13, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 1: Geometric Mean of Minutes per Task with No Minimum Duration Weighted by Task Occasions Observed

PANEL 1: MINUTES PER TASK WITH NO MINIMUM DURATION					
PATIENT POPULATION	CONSULTATIONS OBSERVED	NUMBER OF TASKS WITH NO MINIMUM DURATION INCLUDED IN STUDY	PERFORMANCES OF TASKS WITH NO MINIMUM DURATION OBSERVED PER CHILD	TOTAL TASK OCCASIONS OBSERVED (B*D)	MINUTES PER TASK WITH NO MINIMUM DURATION (CORRECTED FOR RATE AND PACE, IF NECESSARY)
BENIN (CO)	87	19	14.6	1274	1.4
BENIN (SC)	23	19	9.6	222	1.4
BRA_01_C	295	14	9.9	2907	0.5
KEN_03_C	553	32	27.5	15220	0.7
MAR_01_C	397	38	23.7	9400	0.5
SDN_01_C	282	28	16.4	4636	0.8
TZN_01_C	231	21	18.1	4186	0.3
GEOMETRIC MEAN					0.6
PANEL 2: QUALITY OF PROVIDER PERFORMANCE					
PATIENT POPULATION	PERCENT OF PROVIDER CLASSIFICATIONS THAT MATCH THOSE OF SURVEYOR	PERCENT OF CHILDREN CORRECTLY CLASSIFIED	PERCENT OF CHILDREN GIVEN CORRECT SEVERE OR MODERATE CLASSIFICATIONS FOR 3 MAIN SYMPTOMS	PERCENT OF CHILDREN PRESCRIBED UNNECESSARY ANTIBIOTIC	
BENIN (CO)	69.5			NA	
BENIN (SC)	43.0			NA	
BRA_01_C	NA	65.3		7.3	
KEN_03_C	68.9			NA	
MAR_01_C	NA		76.6	23.6	
SDN_01_C	49.1			37.4	
TZN_01_C	NA	63.5		14.0	