Language Matters: Standardised Definitions Can Shape Detection and Research Progress in Congenital Trypanosomiasis

Trypanosoma cruzi, the causative agent of Chagas disease (CD), and *Trypanosoma brucei gambiense* and *Trypanosoma brucei rhodesiense*, the causative agents of Human African trypanosomiasis (HAT) pose considerable health and economic burden in endemic regions. Transmission of these parasites from mothers to children is recognised by the World Health Organization (WHO), yet this transmission route receives significantly less attention than vector borne transmission in the literature and is often absent from elimination strategies, particularly in HAT. Furthermore, terminology describing this transmission route is inconsistent in the literature. Variation in nouns, phrasing, translation and grammar occurs across languages and within individual documents. In under-resourced Neglected Tropical Diseases (NTDs), this inconsistency limits discoverability and prevents rapid synthesis of fragmented evidence, further diminishing the impact of research. Additionally, clinical and policy definitions also diverge. In congenital CD, infants born to infected mothers are considered "at risk" for up to 12 months. By contrast, congenital HAT is typically defined by detection within five days of birth. This study describes the variation in terminology and makes recommendations for a consensus, whilst also describing the history and impact of differing definitions.

A linguistic analysis was conducted on 327 CD publications identified through a narrative review using a backwards citation approach (without search terms). This facilitated inclusion of government and policy documents not retrieved in traditional reviews. The analysis maintained greater objectivity by not relying on predefined search terms. A comparative HAT dataset (n=50) was retrieved following the same method. Publications were grouped by language and systematically examined using a linguistic framework categorising transmission terminology into five categories: Congenital, Maternal, Mother to Child, Vertical and Other. Prevalence, frequency and diacritic-in/sensitive search-term robustness (accent marks on characters do/do not affect results) were quantified.

Terms in the Congenital category had the highest prevalence (92.26% publications) and frequency (76.19% of documented terms). In diacritic-insensitive databases, four search terms were sufficient to retrieve all publications using congenital terminology. Mother to Child terminology appeared in 46.13% publications and required complex proximity searches with diverse orthographies and grammar, limiting reliability as a search term. Vertical terminology was a robust search term but appeared in only 54.49% of publications. Terminological inconsistency was observed within individual WHO documents. Investigating the discrepancy between a 12-month "at risk" period for congenital CD and a five-day diagnosis threshold for congenital HAT found that it reduced opportunities for case finding, without clear biological justification. The HAT definition is impractical in endemic, low resource settings and limits case detection given diagnostic limitations. It is not historically traceable to a biological finding or scientific study. Conversely, the broader CD definition does not imply transmission up to one year after birth, rather it supports testing after maternal antibodies wane, thereby

reducing missed infant diagnoses. Aligning the definition of HAT would establish a definition that is fit for purpose and patient care focussed.

A standardised, WHO-endorsed definition is required for CD and HAT. Adoption of “congenital” terminology and recognition of infants born to infected mothers as “at risk” for up to 12 months, would improve research discoverability, reduce fragmentation of effort, and support a coherent approach to research, policy and clinical care. A sufficient and context-appropriate timeframe for identification and treatment is critical to supporting at-risk infants in both diseases.