Diagnostic Performance of Whole Genome Sequencing x Long-Read Sequencing for Repeat Expansion Disorders: A Systematic Review and Meta-Analysis

AUTORES: Maria Inez Dacoregio[1], Giovanna Giovacchini dos Santos[2], Cainã Rodrigues[3], Heloisa Augusta Castralli[4], Giulia Zoccoli Bueno[5], Rafael Avelar Machado[6], Alexandre Chaves Fernandes[7], Matheus Castro[8], Rachel Honjo[9], Débora Romeo Bertola[10], Chong Ae Kim[11], Paulo Victor Zattar Ribeiro[12]

NOME DAS INSTITUIÇÕES: ICr, Hospital das Clínicas FMUSP[1,4,5,6,7,8,9,10,11], Faculdade de Medicina do ABC[2] , Hospital das Clínicas de Ribeirão Preto USP[12]

INTRODUÇÃO

Repeat expansion disorders represent a major diagnostic challenge, particularly when expansions occur in intronic or non-coding regions. Short-read sequencing technologies, such as Whole Genome Sequencing (WGS), often fail to detect certain pathogenic tandem repeat expansions, especially in genes like RFC1, ATXN3, and CNBP. Long-read sequencing (LRS) technologies, including Oxford Nanopore and PacBio, offer the ability to read long and complex repetitive regions, potentially improving diagnostic yield in this context.

OBJETIVO

To compare the diagnostic performance of WGS and LRS in the detection of repeat expansion disorders through a systematic review and meta-analysis.

METODOLOGIA

We conducted a systematic search of PubMed and Scopus up to June 2025. Eligible studies included original research using WGS or LRS in individuals clinically suspected of having repeat expansion disorders. Outcomes included diagnostic yield (proportion of molecular diagnoses) and odds ratio (OR) comparing technologies. Meta-analyses were performed using random-effects models (DerSimonian–Laird method) with the meta and ggplot2 packages in R.

RESULTADOS E DISCUSSÃO

A total of 14 studies involving 30,500 individuals were included. The pooled diagnostic yield was 8% for WGS (95% CI: 5–12%; I^2 = 77%) and 53% for LRS (95% CI: 35–71%; I^2 = 98%), favoring long-read technologies. In five gene-matched studies, the pooled odds ratio for molecular diagnosis using LRS over WGS was 19.53 (95% CI: 3.92–97.31; I^2 = 84%), with individual ORs ranging from 4.2 to 32.5. This superiority was consistently observed in genes such as RFC1, ATXN3, and SAMD12. Log-scale forest plots highlighted the magnitude and consistency of this diagnostic advantage.

These results confirm that LRS technologies substantially outperform WGS in the detection of pathogenic repeat expansions. While WGS is a valuable genome-wide tool, it remains limited for resolving complex tandem repeat regions. In contrast, LRS enables direct, base-level characterization of expansions, leading to greater diagnostic confidence. Despite inter-study heterogeneity, the direction and strength of the effect consistently favor LRS.

CONCLUSÃO

Long-read sequencing provides significantly higher diagnostic yield than short-read WGS for repeat expansion disorders. Its clinical implementation may reduce the diagnostic odyssey and improve precision care for individuals with suspected expansion-mediated diseases.

Figure 1: Diagnostic yield of WGS and LRS for repeat-expansion disorders: overall meta-analysis of proportions

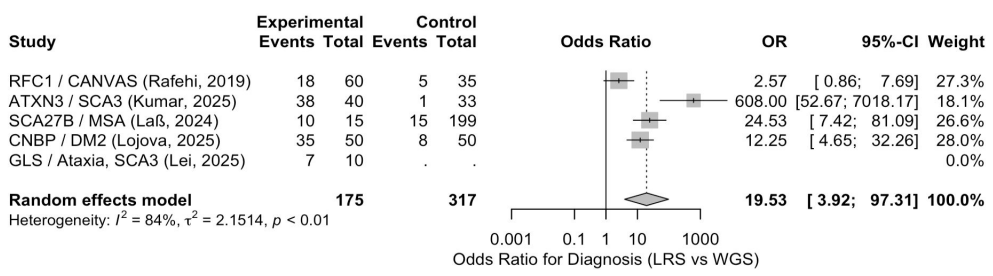
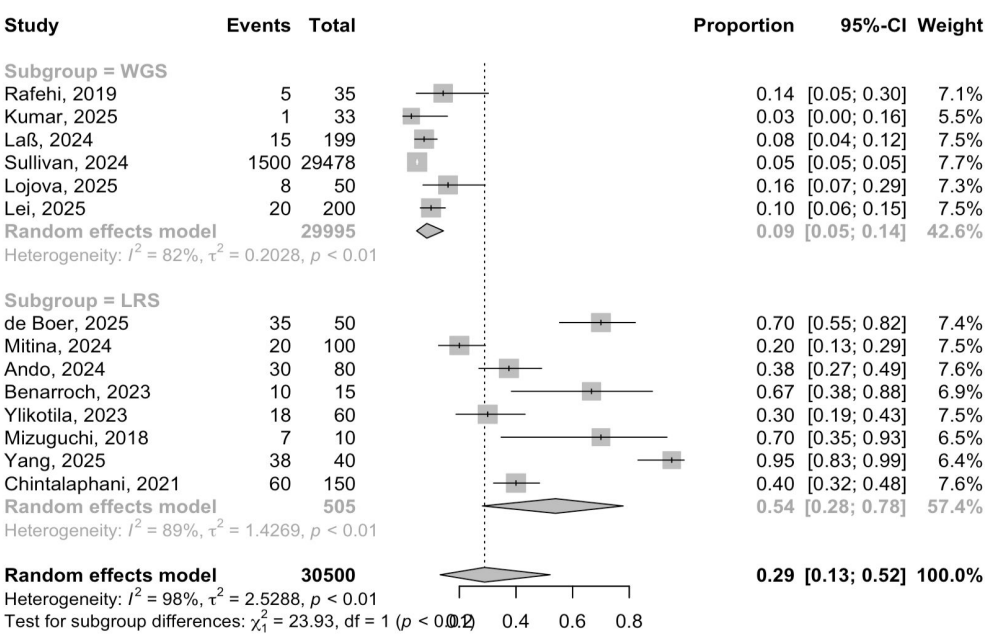


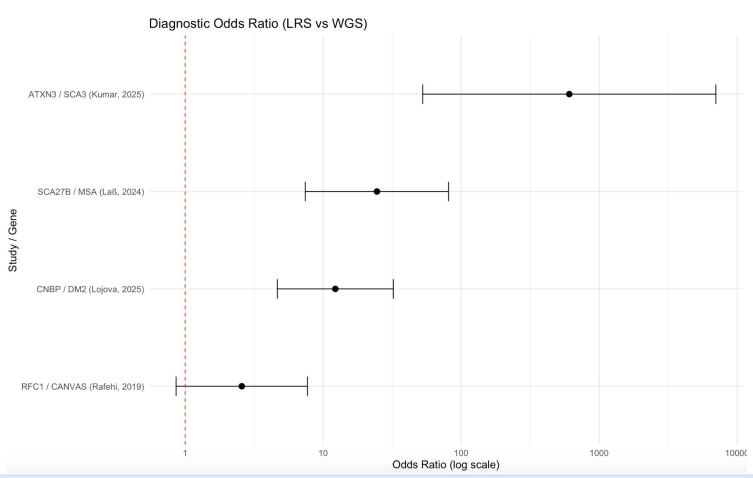
Figure 2: Subgroup meta-analysis of diagnostic yield by sequencing technology (WGS vs LRS)



Gene / Condition	WGS Study (Author / Year)	WGS N total	WGS N positive
RFC1 / CANVAS	Rafehi, 2019	35	5
ATXN3 / SCA3	Kumar, 2025	33	1
SCA27B / MSA	Laß, 2024	199	15
CNBP / DM2	Lojova, 2025	50	8

Gene / Condition	LRS Study (Author / Year)	LRS N total	LRS N positive
RFC1 / CANVAS	Ylikotila, 2023	60	18
ATXN3 / SCA3	Yang, 2025	40	38
SCA27B / MSA	Benarroch, 2023	15	10
CNBP / DM2	de Boer, 2025	50	35
SAMD12 / FCMTE	Mizuguchi, 2018	10	7

Author / Year	Sequencing Type	Gene / Condition	N total	N positive	Diagnostic yield (%)
Kumar, 2025	WGS	ATXN3 / SCA3	33	1	3
Rafehi, 2019	WGS	RFC1 / CANVAS	35	5	14,3
Laß, 2024	WGS	SCA27B / MSA	199	15	7,5
Sullivan, 2024	WGS	RFC1 / Neurological	29478	1500	5,1
Lojova, 2025	WGS	CNBP / DM2	50	8	16
de Boer, 2025	LRS	Multiple STRs	50	35	70
Yang, 2025	LRS	Multiple STRs	40	38	95
Mizuguchi, 2018	LRS	SAMD12 / FCMTE	10	7	70
Benarroch, 2023	LRS	OPDM	15	10	66,7



REFERÊNCIAS

1.Kaplun, Ludmila et al. “ONT in Clinical Diagnostics of Repeat Expansion Disorders: Detection and Reporting Challenges.” International journal of molecular sciences vol. 26,6 2725. 18 Mar. 2025, doi:10.3390/ijms26062725. 2.International journal of molecular sciences vol. 26,6 2725. 18 Mar. 2025, doi:10.3390/ijms26062725

AGRADECIMENTOS: Agradeço a Deus, minha família, ao grupo de pesquisa de meta análises com grande honra ao Prof. Dr Rhanderson Cardoso, aos meus colegas de residência e aos meus chefes e professores da Residência de Genética Médica do ICr HCFMUSP.